Advances in genetics and molecular breeding of three legume crops of semi-arid tropics using next-generation sequencing and high-throughput genotyping technologies

Rajeev K Varshney[1,2,*], Himabindu Kudapa[1], Manish Roorkiwal[1], Mahendar Thudi[1], Manish K Pandey[1], Rachit K Saxena[1], Siva K Chamarthi[1], Murali Mohan S[1], Nalini Mallikarjuna[1], Hari Upadhyaya[1], Pooran M Gaur[1], L Krishnamurthy[1], KB Saxena[1], Shyam N Nigam[1] and Suresh Pande[1]

[1]*International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India*

[2]*CGIAR Generation Challenge Programme (GCP), c/o CIMMYT, 06600, Mexico, DF, Mexico*

**Corresponding author (Fax, 0091-40-30713074/75; Email, r.k.varshney@cgiar.org)*

Molecular markers are the most powerful genomic tools to increase the efficiency and precision of breeding practices for crop improvement. Progress in the development of genomic resources in the leading legume crops of the semi-arid tropics (SAT), namely, chickpea (*Cicer arietinum*), pigeonpea (*Cajanus cajan*) and groundnut (*Arachis hypogaea*), as compared to other crop species like cereals, has been very slow. With the advances in next-generation sequencing (NGS) and high-throughput (HTP) genotyping methods, there is a shift in development of genomic resources including molecular markers in these crops. For instance, 2,000 to 3,000 novel simple sequence repeats (SSR) markers have been developed each for chickpea, pigeonpea and groundnut. Based on Sanger, 454/FLX and Illumina transcript reads, transcriptome assemblies have been developed for chickpea (44,845 transcript assembly contigs, or TACs) and pigeonpea (21,434 TACs). Illumina sequencing of some parental genotypes of mapping populations has resulted in the development of 120 million reads for chickpea and 128.9 million reads for pigeonpea. Alignment of these Illumina reads with respective transcriptome assemblies have provided >10,000 SNPs each in chickpea and pigeonpea. A variety of SNP genotyping platforms including GoldenGate, VeraCode and Competitive Allele Specific PCR (KASPar) assays have been developed in chickpea and pigeonpea. By using above resources, the first-generation or comprehensive genetic maps have been developed in the three legume species mentioned above. Analysis of phenotyping data together with genotyping data has provided candidate markers for drought-tolerance-related root traits in chickpea, resistance to foliar diseases in groundnut and sterility mosaic disease (SMD) and fertility restoration in pigeonpea. Together with these trait-associated markers along with those already available, molecular breeding programmes have been initiated for enhancing drought tolerance, resistance to fusarium wilt and ascochyta blight in chickpea and resistance to foliar diseases in groundnut. These trait-associated robust markers along with other genomic resources including genetic maps and genomic resources will certainly accelerate crop improvement programmes in the SAT legumes.

Keywords. Chickpea; genomic resource; genotyping; groundnut; legume; pigeonpea; sequencing

Abbreviations used: AB, ascochyta blight; AFLP, amplified fragment length polymorphism; BAC, bacterial artificial chromosome; BES, BAC- end derived sequence; CAPS, cleaved amplified polymorphic sequences; CISR, conserved intron spanning region; COS, conserved orthologous sequence; DArT, diversity array technology; FW, Fusarium wilt; GMM, genic molecular marker; GS, Genomic selection; HTP, high-throughput; LLS, late leaf spot; MABC, marker-assisted backcrossing; MAS, marker-assisted selection; NGS, next-generation sequencing; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; RIL, recombinant inbred line; SAT, semi-arid tropics; SCMR, SPAD chlorophyll meter readings; SLA, specific leaf area; SMD, sterility mosaic disease; SNP, single nucleotide polymorphism; SPAD, soil plant analytical development; SSR, simple sequence repeats; TAC, transcript assembly contig; TE, transpiration efficiency; TUS, tentative unique sequences

Published online: 15 October 2012

[Varshney RK, Kudapa H, Roorkiwal M, Thudi M, Pandey MK, Saxena RK, Chamarthi SK, Mohan SM, Mallikarjuna N, Upadhyaya H, Gaur PM, Krishnamurthy L, Saxena KB, Nigam SN and Pande S 2012 Advances in genetics and molecular breeding of three legume crops of semi-arid tropics using next-generation sequencing and high-throughput genotyping technologies. *J. Biosci.* **37** 811–820] **DOI** 10.1007/s12038-012-9228-0

1. Introduction

Legumes include a number of important crops, namely, soybean (*Glycine max*), groundnut (*Arachis hypogaea*), cowpea (*Vigna unguiculata*), common bean (*Phaseolus vulgaris*), chickpea (*Cicer arietinum*), pigeonpea (*Cajanus cajan*), lentil (*Lens culinaris*), pea (*Pisum sativum*), mungbean (*Vigna radiata*), etc. These legume crops fall into two Papilionoid clades, namely, Galegoid and Phaseoloid, which are often referred to as cool season and warm (or tropical) season legumes, respectively (Lewis *et al.* 2005). The global production of grain and forage legumes is about 300 million metric tons (Vance *et al.* 2000) which are grown on about 190 million hectares. Legumes are a rich source of proteins, vitamins, minerals and dietary fibre (Duranti and Gius 1997; Grusak 2002).

Several food legume crops are grown in semi-arid tropics (SAT) of Africa, Asia and South America. In these areas, the legume crops are exposed to various biotic and abiotic stresses. As a result in SAT areas average crop productivity of the majority of legume crops like chickpea, pigeonpea and groundnut is around 1 ton/ha (*www.fao.org*) or even less. Crop productivity of these legume crops can be enhanced through the use of biotechnological tools in the breeding programmes. Marker trait associations are a prerequisite for marker-assisted selection (MAS) (Varshney *et al.* 2005a, 2009a). Marker trait association for a number of traits in all major crops have now become available due to the accessibility of an array of molecular markers and dense molecular genetic maps, and so MAS has become routine in breeding programmes of several major crop species (Varshney *et al.* 2006; Kulwal *et al.* 2011). However, majority of the legume crops, unlike the model legumes like *Medicago* (*Medicago truncatula*) and *Lotus* (*Lotus japonicus*) or industrial crops including soybean (*Glycine max*) remained untouched with genomics revolution (Wilson *et al.* 2004; Varshney *et al.* 2010). ICRISAT together with its partners, during the last few years, have developed significant amount of genomic resources that have already started to make an impact on trait mapping and molecular breeding in the above-mentioned legume crops. This article reviews the progress on development and application of genomics resources in accelerating genomics research and breeding applications in the SAT legume crops.

2. Genetic resources and their utilization in modern breeding

Analysis of molecular diversity and pedigree information of the cultivars developed through classical breeding approaches indicated a narrow genetic base in cultivated genepools of chickpea, pigeonpea and groundnut. However, a large number of accessions for each of the three legume species are present in several genebanks of the world (Bohra *et al.* 2011a; Pandey *et al.* 2011; Upadhyaya *et al.* 2011). The ICRISAT genebank stores the largest number of accessions for each of the above-mentioned legume crops. In order to enhance the use of available genetic resources in crop breeding, the concept of core collection (Brown 1989), mini core collection (Upadhyaya and Ortiz 2001) and core reference set (Glaszmann *et al.* 2010) has been followed. These collections are now available in all these three legume species (see recent reviews by Bohra *et al.* 2011a; Pandey *et al.* 2011; Upadhyaya *et al.* 2011).

Development of mapping populations using identified genetic resources for trait mapping and marker-assisted breeding is the subsequent important task. Ideally, the suitable genotypes for developing the mapping population should be genetically diverse and with contrast phenotype for the trait of interest (Saxena *et al.* 2010a). However, due to narrow genetic diversity in the cultivated genepools of the SAT legumes, identification of such genotypes is quite challenging. Nevertheless, by using conventional approaches as well as by using molecular diversity information, several mapping populations have been developed at ICRISAT and its partner institutes. Although these mapping populations are good for marker trait association, they may not be suitable for developing dense genetic maps. Therefore, inter-specific mapping populations have also been developed and used for developing genetic maps with higher marker density in chickpea and pigeonpea (table 1). However, in groundnut, genetic maps are being developed using inter-specific mapping population derived from the cross TAG 24 × ICGV 86031, and consensus map is being developed using several mapping populations.

3. Genomic resources for SAT legumes

In the past, for genetic diversity analysis, a range of molecular markers such as restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs) and

Table 1. Development of reference genetic maps for SAT legumes

Crop	Mapping population	No. of progenies/ type of mapping population	No. of mapped markers	Reference
Chickpea	ICC 4958 × PI 489777	131 RILs	1291	Nayak *et al.* 2010 Thudi *et al.* 2011
Groundnut	TAG 24 × ICGV 86031	318 RILs	191	Varshney *et al.* 2009d Ravi *et al.* 2011
Pigeonpea	ICP 28 × ICPW 94	79 F_2 lines	239	Bohra *et al.* 2011b

simple sequence repeats (SSRs) have been used. However, SSR and single nucleotide polymorphism (SNP) markers have become the markers of choice for genetic analysis and breeding applications in the SAT legume crops (Varshney *et al.* 2007). While SSR markers have the advantages of being multi-allelic and co-dominant (Gupta and Varshney 2000), SNP markers offer high-throughput and cost-effective genotyping options. Another high-throughput marker system, is diversity array technology (DArT), which became popular in many other crop species since no sequence information is needed for developing these markers (Killian *et al.* 2005).

The last 5 years have witnessed significant progress in the area of development of genomic resources in these SAT legume crops that have made them 'genomic resource rich' crops from so-called 'orphan crops'. This has been possible due to collaborative and coordinated efforts of the legume community and financial support from a number of organizations including CGIAR Generation Challenge Programme/The Bill & Melinda Gates Foundation, Indian Council of Agricultural Research (ICAR), Department of Biotechnology (DBT) of Government of India, US National Science Foundation (NSF), etc.

3.1 *SSR markers*

Until recently, very few SSR markers were available in three SAT legume species. During the last few years, a large number of SSR markers in each of three legumes have been developed by using following approaches individually or in combination: (a) constructing and sequencing of SSR-enriched genomic DNA libraries, (b) sequencing and mining the BAC (bacterial artificial chromosome)-end sequences (BES) for SSRs, and (c) mining the transcript sequences generated by either Sanger sequencing or next-generation sequencing (NGS) approaches such as 454/FLX sequencing (figure 1). Using these approaches independently or in combination, about 2,000 novel SSR markers have been developed in chickpea (Nayak *et al.* 2010; Thudi *et al.* 2011), 3,200 in pigeonpea (Saxena *et al.* 2010b; Bohra *et al.* 2011b; Dutta *et al.* 2011) (table 2) and about 2,500 in groundnut (Mace *et al.* 2007; Cuc *et al.* 2008; Gautami *et al.* 2009; Pandey *et al.* 2012).

3.2 *DArT markers*

The DArT marker system has been widely used for constructing genetic maps and diversity analysis in *Triticeae* species (Wenzl *et al.* 2006; Neumann *et al.* 2011; Roy *et al.* 2011; Varshney *et al.* 2012a). In the case of legumes, the first-generation array comprising 6,144 clones was developed in pigeonpea (Yang *et al.* 2011). Recently, DArT arrays have become available in common bean also (Briñez *et al.* 2011). With DArT Pty Ltd, ICRISAT has developed DArT arrays comprising 15,360 clones in each of the three species. Use of these DArT arrays, as expected, showed a narrow genetic diversity in the elite genepool as compared to landraces and wild species. The parental genotypes of mapping populations including intra-specific mapping populations in chickpea and pigeonpea, when screened with the available DArT arrays, showed 35% and 9% polymorphism, respectively. In summary, it seems that DArT markers are not cost-effective or attractive marker system for detecting polymorphism in cultivated germplasm of the SAT legume crops. However, DArT markers may prove useful for introgression of segments from alien species to the elite varieties of the legume crops. For instance, in pigeonpea by using 1,225 DArT markers in the cross between *C. platycarpus* and *C. cajan*, 2–5% *C. platycarpus* genome carrying genes for disease and insect resistance was observed (Mallikarjuna *et al.* 2011).

3.3 *Transcript assembly and SNP markers*

NGS technologies offer the ability to produce huge sequence data sets at relatively low cost in less time. These technologies are based on a combination of template preparation, sequencing, alignment and assembly of the genome. Although a number of second-generation or third-generation sequencing technologies have become available (Thudi *et al.* 2012), most commonly used sequencing technologies are 454 (454 Life Sciences, *http://www.my454.com/*), SOLiD (Applied Biosystems, *http://www3.appliedbiosystems.com*) and Illumina (Illumina Inc., *http://www.my454.com/*) (Varshney *et al.* 2009b).

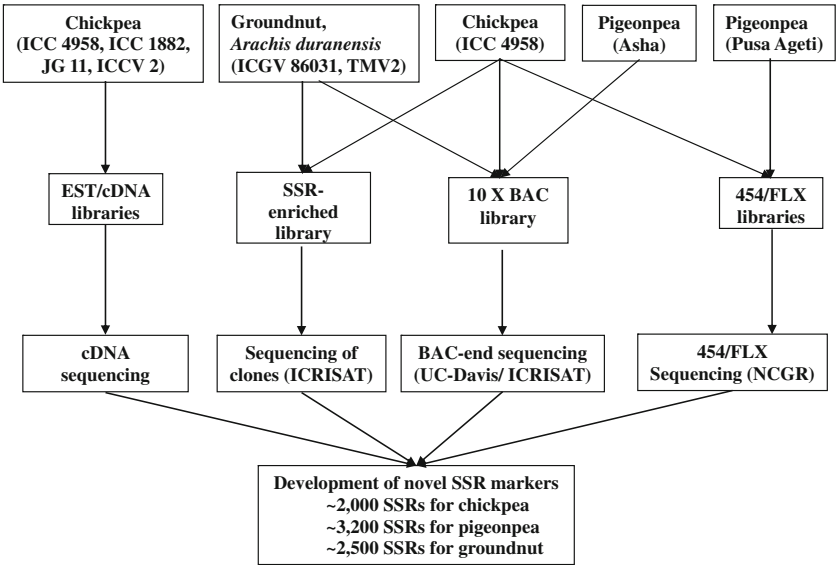

Figure 1. Schematic representation of microsatellite development and microsatellite enriched libraries, BAC-end sequences and transcriptomic resources developed through Sanger and next-generation sequencing technologies.

Two NGS technologies namely 454 and Illumina together with Sanger sequencing technology have been used to characterize the transcriptomes of chickpea and pigeonpea. For instance, 20,162 and 9,888 ESTs were developed for chickpea and pigeonpea using Sanger sequencing technology on drought- and salinity-challenged cDNA libraries for chickpea (Varshney *et al.* 2009c) and Fusarium wilt (FW) and sterility mosaic disease (SMD)-challenged cDNA libraries for pigeonpea (Raju *et al.* 2010). By analysing these ESTs together with the then available ESTs in public domain, 9,569 unigenes for chickpea and 5,085 for pigeonpea were identified. In order to improve these transcriptomic resources, 454/FLX sequencing was undertaken on normalized and pooled RNA samples collected from >20 tissues representing different developmental stages of the plant. As a result, 435,018 transcript reads for chickpea and 494,353 transcript reads for pigeonpea have been generated (Dubey *et al.* 2011; Hiremath *et al.* 2011). Cluster analysis of these transcript reads with Sanger ESTs generated at ICRISAT as well as those available in public domain provided transcript assembly of chickpea (Ca TA) with 103,215 tentative unique sequences (TUSs) and pigeonpea (Cc TA v1) with 127,754 TUSs (Dubey *et al.* 2011; Hiremath *et al.* 2011). In the case of pigeonpea TA, 494,353 454/FLX transcript reads generated from Asha genotype and 128.9 million Illumina reads generated from 12 genotypes were analysed together with 18,353 Sanger ESTs and 1.696 million 454/FLX transcript reads (Dutta *et al.* 2011) with improved algorithms. As a result, an improved TA in pigeonpea referred as Cc TA v2 comprising 21,434 contigs has been developed (Kudapa *et al.* 2012).

Table 2. Transcriptomic resources and molecular markers developed at ICRISAT using next generation sequencing and high-throughput genotyping technologies

Resource	Chickpea	Pigeonpea
SSRs	~2,000	3,200
SNPs	9,000	10,000
GoldenGate assays	768 SNPs	768 SNPs
KASPar assays	2,005 SNPs	1,616 SNPs
Sanger ESTs	~30,000	~20,000
454 /FLX reads	435,018	494,353
TUSs	103,215	21,432
Illumina reads (million reads)	>108 (4 parents)	>160 (16 parents)

Three approaches were used for identification of SNPs. In the first approach, Illumnia sequencing was carried out on parental genotypes of mapping populations of chickpea and pigeonpea. RNA sequencing of 4 chickpea and 12 pigeonpea genotypes has resulted ca. 120 million reads for chickpea (Hiremath *et al.* 2011) and 128.9 million reads for pigeonpea (Kudapa *et al.* 2012). Alignment of these short reads onto TAs of respective species, as mentioned above, has provided a large number (tens of thousands) of SNPs in each of these species (table 2). The second approach of allele-specific sequencing of parental genotypes of the reference mapping populations of chickpea and pigeonpea using conserved orthologous sequence (COS) markers has provided 768 SNPs each for chickpea and pigeonpea, respectively. In the third approach, allele-specific sequencing for 220 candidate genes was undertaken on 2–20 chickpea genotypes and

1,893 SNPs were identified (Gujaria *et al.* 2011). In brief, a large number of SNPs have become available for chickpea and pigeonpea and high- as well as low- throughput and cost-effective genotyping platforms have been developed.

4. Cost-effective SNP genotyping platforms

Use of a particular marker system for genetics research and breeding application depends on the throughput and cost of the marker assays. Although a large number of SNPs have become available in chickpea (Hiremath *et al.* 2011) and pigeonpea (Dubey *et al.* 2011), different applications need a varying level of throughput of SNP genotyping. While GoldenGate assays with a possibility to undertake genotyping of 768 SNPs have been developed in both species, VeraCode assays for genotyping 96 SNPs in chickpea and 48 SNPs in pigeonpea have also been developed (figure 2). For several breeding applications only a few SNP markers are required as in the foreground selection during MAS experiments. In those cases, GoldenGate or VeraCode assays may not be cost-effective. Therefore, KASPar assays have been developed for 2,005 SNPs in chickpea and 1,616 SNPs in pigeonpea (table 2 and figure 2). For example, very few markers are required for genotyping large-scale segregating populations during marker-assisted selection (MAS) programmes. In such cases KASPar assays from KBiosciences (*www.kbioscience.co.uk*) seem to be worthwhile. In addition, 279 CAPS (cleaved amplified polymorphic sequences) and 121 CISR (conserved intron spanning region) markers have also been developed for SNP genotyping in the case of chickpea (Gujaria *et al.* 2011). For pigeonpea 6,284 CISR markers were identified (Kudapa *et al.* 2012).

5. Genetic and transcript maps

Available molecular markers (SSRs, SNPs and DArTs) have been used for constructing genetic maps in the above-mentioned legume crops (Varshney *et al.* 2009c). For instance, the first SSR-based genetic linkage map was developed for groundnut based on TAG 24 × ICGV 86031 (RIL-1) recombinant inbred line (RIL) mapping population. This map was further saturated up to 191 SSR loci (Ravi *et al.* 2011) (table 1) followed by construction of two more new genetic maps using the RILs ICGS 76 × CSMG 84-1 (RIL-2; 119 SSR loci) and ICGS 44 × ICGS 76 (RIL-3; 82 SSR loci) segregating for drought-tolerance-related traits. Together with these maps and the reference map with 191 SSR loci based on RIL-1, a consensus map was constructed with 293 SSR loci spanning 2,840.8 cM (Gautami *et al.* 2011). Similarly, two new genetic maps based on mapping population derived from crosses RILs TAG 24 × GPBD 4 (RIL-4; 188 SSR loci) and TG 26 × GPBD 4 (RIL-5: 181 SSR loci) segregating for foliar disease resistance were constructed and were used for development of a consensus map with 225 SSR loci with a total map distance of 1152.9 cM (Sujay *et al.* 2011).

In the case of chickpea, genetic maps were available earlier based on the inter-specific mapping population (ICC 4958 × PI 489777). These maps were developed using anonymous markers like RAPD and AFLP or with smaller number of SSR markers. With the intent of extending the genetic map, and enhancing the number of easily scorable markers, a high-density chickpea genetic map with 1,291 loci has been developed by Thudi *et al.* (2011). This map has 157 novel SSR markers developed from BES-SSRs, 621

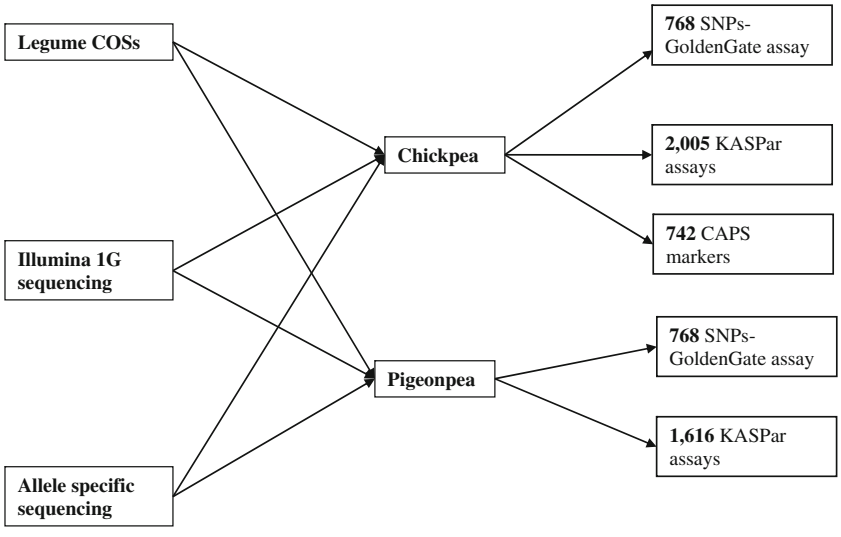

Figure 2. Development of gene-based SNP markers for chickpea and pigeonpea using next-generation sequencing and high-throughput genotyping technologies.

DArT loci, 145 EST-derived molecular markers and 368 legacy markers published earlier, with eight linkage groups and spanning a distance of 845.56 cM. The number of markers on each linkage group ranged from 68 (LG 8) to 218 (LG 3) with an average inter-marker distance of 0.65 cM. As molecular markers are being developed now from gene sequences/transcripts, the genetic maps integrated with such markers are also referred as 'transcript maps'. In the case of chickpea, a transcript map comprising 126 genic molecular marker (GMM) loci has been developed. These marker loci included 53 CAPS-SNPs, 55 EST-SSRs and 18 CISR. Recently, 1,189 KASPar-assay-based SNPs have also been added to this transcript map and the updated map has 1,328 gene-based marker loci (Hiremath *et al.* unpublished).

In the case of pigeonpea, efforts have been made only recently to develop genetic maps. Based on an inter-specific mapping population (ICP 28 × ICPW 94), DArT and bacterial artificial chromosome (BAC)-end derived sequence (BES) SSR markers were used for developing genetic maps. A total of 122 DArT markers were used to generate maternal genetic linkage map (270.0 cM) and 172 DArT markers were used to generate a paternal genetic linkage map. The average marker distance is 2.2 cM and 2.6 cM, per marker in maternal and paternal linkage maps, respectively (Yang *et al.* 2011). Another BES-SSR-based genetic map was developed by Bohra *et al.* (2011b) using 239 BES-SSR markers covering 930.90 cM with an average of 21 markers per linkage groups and an average marker distance of 3.8 cM (table 1).

5.1 *Trait mapping*

Yield in SAT legumes is mainly affected by various abiotic stresses like drought and biotic stresses such as FW and SMD in pigeonpea and foliar diseases in groundnut. To overcome the constraints related to production in SAT legumes using genomics-assisted breeding, identification/discovery of marker trait association between the trait of interest and a genetic marker is an important and starting point to work for crop improvement. Because of paucity of markers and non-availability of genetic maps, QTL mapping in SAT legumes have been very slow. However, during the last 5 years, ICRISAT in collaboration with its partners have made some progress towards QTL mapping for several production constraints.

Drought tolerance is a complex trait and several component traits include root traits, transpiration efficiency (TE), and its surrogates such as carbon isotope discrimination (^{13}C), specific leaf area (SLA) or soil plant analytical development (SPAD) chlorophyll meter readings (SCMR). For molecular mapping, two intra-specific mapping populations of chickpea (ICC 4958 × ICC 1882 and ICC 283 × ICC 8261) have been developed at ICRISAT using drought-tolerant genotypes (ICC 4958 and ICC 8261) and drought sensitive genotypes (ICC 1882 and ICC 283) (Chamarthi *et al.* 2011). These populations have been genotyped and phenotyped extensively. Based on genotyping data, two genetic maps comprising 240 and 170 SSR loci have been developed for ICC 4958 × ICC 1882 and ICC 283 × ICC 8261 mapping populations, respectively. QTL analysis based on these genetic maps and phenotyping data obtained for 11 root traits in 2 years provided a genomic region that contains QTLs for several root-related traits contributing ~30% phenotypic variation. Incidentally, this genomic region also has QTLs for other drought-tolerance-related traits like C-isotope discrimination, yield and harvest index in rain-fed conditions. Therefore, this genomic region was targeted for introgression in elite chickpea lines for enhancing drought tolerance using marker-assisted backcrossing approach.

In the case of groundnut, three mapping populations, namely, TAG 24 × ICGV 86031, ICGS 44 ×ICGS 76 and ICGS 76 × CSMG 84-1, were targeted for mapping drought tolerance. These populations were phenotyped for transpiration, transpiration efficiency, biomass, specific leaf area, pod weight, total dry matter, SPAD chlorophyll meter reading, total dry weight, shoot dry weight and harvest index traits in 2–3 seasons during 2004, 2008 and 2009. In parallel, after screening 3,215 markers, genetic maps were developed for all the three mapping populations that comprise 82 (ICGS 44 × ICGS 76) to 191 (TAG 24 × ICGV 86031) marker loci. Detailed analysis using the phenotyping data and genotyping as mentioned above with different programs like QTL Cartographer, QTL Network and Genotype Matrix Mapping identified 153 main effect and 25 epistatic QTLs for drought-tolerance-related traits (Varshney *et al.* 2009c; Gautami *et al.* 2011; Ravi *et al.* 2011). Majority of these QTLs contribute relatively low phenotypic variation. Molecular breeding approaches like marker-assisted backcrossing (MABC) may not be useful for introgressing drought tolerance in elite varieties, and some other approaches like marker-assisted recurrent selection (MARS) and genomic selection (GS) may be preferred approaches for such purpose. In addition to drought, late leaf spot (LLS) and rust are two main foliar diseases of groundnut. ICRISAT in collaboration with the University of Agricultural Sciences–Dharwad (UAS-D) have identified markers associated with these two diseases. In this regard, two RIL populations, namely, TAG 24 × GPBD 4 and TG 26 × GPBD 4 comprising 268 and 146 lines, respectively, were extensively phenotyped for rust and LLS resistance for 7–8 seasons (2004–2010) at UAS-D. After screening a total of 3,097 SSR markers among parental genotypes of these two populations, a total of 209 polymorphic markers were identified for each of the two populations at ICRISAT. Genetic linkage maps were constructed for TAG 24 × GPBD 4 (188 loci) and TG 26 × GPBD 4 (181 loci). By using genotyping and phenotyping data on these populations, a total of 28 QTLs for LLS and 13 QTLs for rust explaining

10.07 to 67.8% and 2.54 to 82.96% phenotypic variation, respectively, were detected (Khedikar *et al.* 2010; Sujay *et al.* 2011). Associated markers (IPAHM 103, GM2079, GM2301 and GM1536) linked with rust QTL in TAG 24 × GPBD 4 were validated in another mapping population (TG 26 × GPBD 4) and in a set of resistant/susceptible breeding lines. Furthermore, an attempt was also made to identify linked markers for important nutritional traits using the genotyping and phenotyping data generated on TG26 × GPBD4. The phenotyping data was generated for protein content, oil content and oil quality at UAS-D, while genotyping with 53 polymorphic markers was generated at ICRISAT. QTL analysis detected a total of seven QTLs for protein content (2.54–9.78%), eight QTLs for oil content (1.5.10.2%) and six common QTLs for oleic and linoleic acid (3.3–9.7%) (Sarvamangala *et al.* 2011). These markers were then used for introgression of this major QTL for rust resistant in three elite cultivars using MABC approach.

In pigeonpea, QTL analysis provided six QTLs for SMD using ICP 8863 × ICPL 20097 and TTB 7 × ICP 7035 mapping populations. Of these one QTL (qSMD4) explaining 24.72% of phenotypic variance was identified on LG 7 (Gnanesh *et al.* 2011). Furthermore, several QTLs explaining up to 24% phenotypic variation have been identified for fertility restoration (*Rf*) in pigeonpea (Bohra *et al.* unpublished).

6. Molecular breeding

After identification of QTLs/genes responsible for trait of interest, the next step is to use this information in crop improvement. However, for successful introgression of QTLs in elite breeding materials or varieties, the targeted QTLs should be major QTLs that contribute >20% phenotypic variation. Amongst the three legume crops mentioned in this article, molecular breeding efforts have been initiated in chickpea and groundnut.

In the case of chickpea, molecular breeding programmes are underway for enhancing drought tolerance and disease resistance. For drought tolerance, nine leading chickpea varieties namely JG 11, Chefe, KAK2, Arerti, ICCV 10, ICCV 95423, ICCV 97105, Ejere and DCP92-3 have been targeted by ICRISAT and its partners Egerton University (Kenya), Ethiopian Institute of Agricultural Research (EIAR, Ethiopia), Lake Zone Agricultural Research Development Institute (LZARDI, Tanzania), Indian Institute of Pulse Research (IIPR, India) and Indian Agricultural Research Institute (IARI, New Delhi). In the case of JG 11, Chefe and KAK 2 varieties, BC_3F_3 seeds have been produced. Preliminary analysis of $BC_3F_{3:4}$ lines for root traits gave encouraging results and agronomic performance is being evaluated under irrigated and rain-fed conditions at different locations.

Similarly, in another molecular breeding network project of DBT, lead by ICRISAT along with Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV, Jabalpur), Indian Institute of Pulses Research (IIPR, Kanpur), Mahatma Phule Krishi Vidyapeeth (MPKV, Rahuri) and Agricultural Research Station (ARS, Gulbarga), efforts have been initiated to introgress resistance to fusarium wilt (FW) and ascochyta blight (AB) in elite chickpea cultivars (C 214, JG 74, Pusa 256, Phule G12 and Annigeri-1) from different agro-climatic zones through MABC. In the case of variety C 214 also, BC_3F_3 lines have been generated, while in the case of varieties JG 74, Pusa 256, Phule G12 and Annigeri-1, respectively, BC_2F_1 and BC_1F_1 lines have been generated.

In the case of groundnut, efforts are in progress to introgress resistance to rust disease in varieties ICGV 91114, JL 24 and TAG 24 at ICRISAT. For this purpose, a major QTL contributing 82.96% phenotypic variation has been targeted. Four SSR markers present in the QTL region (IPAHM 103, GM2079, GM2301 and GM1536) for rust resistance are being used for foreground selection. As a result of MABC, 76 homozygous BC_3F_2 and 158 BC_2F_3 lines have been generated in the genetic background of ICGV 91114, JL 24 and TAG 24 (Pandey *et al.* 2012). Initial screening of these lines for rust has identified several promising lines that showed remarkable reduction in disease spread.

7. Prospects of genomics-assisted breeding in SAT legumes

It is evident from the above discussions that as a result of concerted and coordinated efforts of several partners together with advances in genomics like NGS and HTPG technologies, the so-called 'orphan legume crops' have become 'genomic resources rich crops' now (Varshney *et al.* 2009b). In the case of pigeonpea, even the genome sequence has become available (Varshney *et al.* 2012b). Similar efforts are underway in chickpea and groundnut. While molecular markers have already become available for several important traits like FW and AB resistance and drought tolerance in chickpea, foliar diseases resistance in groundnut, and SMD and *Rf* in pigeonpea. Efforts have already been initiated to use the linked markers for molecular breeding applications for enhancing drought tolerance in chickpea and disease resistance in both chickpea and groundnut. It is anticipated that availability of large-scale genomic resources and cost-effective genotyping platforms as well as possibilities of outsourcing of genotyping work will accelerate trait mapping for other important traits in the legume crops. This will eventually enhance adoption of molecular breeding for increasing crop productivity in these legume crops.

Acknowledgements

We acknowledge the collaboration and support of several colleagues and collaborators from ICRISAT, University of California, Davis (UC-Davis, USA), National Center for

Genome Research (NCGR, USA), J. Craig Venter Institute (JCVI, USA), University of Frankfurt, Germany, Diversity Arrays Technology (DArT, Australia) Pty Ltd, National Research Center on Plant Biotechnology (NRCPB, India), Indian Institute of Pulse Research (IIPR, India), National Institute for Plant Genome Research (NIPGR, India), Directorate of Groundnut Research (DGR, India), University of Agricultural Sciences, Dharwad (UAS-D, India), Mahatma Phule Krishi Vidyapeeth, Rahuri (MPKV-India), Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV, India), ICAR-Agricultural Research Station, Gulbarga (ARS-Gulbarga, India), Egerton University (EU, Kenya), Ethiopian Institute of Agricultural Research (EIAR, Ethiopia) and Lake Zone Agricultural Research Development Institute (LZARDI, Tanzania) that have made it possible to develop and apply genomic resources for trait mapping and molecular breeding mentioned in the article. The authors are thankful to several funding agencies like CGIAR Generation Challenge Programme, Tropical Legume I Initiative of Bill and Melinda Gates Foundation, Indian Council of Agricultural Research (ICAR), Department of Biotechnology (DBT) and Department of Science and Technology (DST) and Indo-German Science and Technology Centre for funding genomics and breeding projects in the mentioned legume crops.

References

Bohra A, Dubey A, Saxena RK, Penmetsa RV, Poornima KN, Kumar N, Farmer AD, Srivani G, *et al.* 2011b 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

Bohra A, Mallikarjuna N, Saxena K, Upadhyaya H, Vales MI and Varshney RK 2011a Harnessing the potential of crop wild relatives through genomics tools for pigeonpea improvement. *J. Plant Biol.* **37** 1–16

Briñez B, Blair MW, Kilian A, Carbonell SAM, Chiorato AF and Rubiano LB 2011 A whole genome DArT assay to assess germplasm collection diversity in common beans. *Mol. Breed.* doi: 10.1007/s11032-011-9609-3

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

Chamarthi SK, Kumar A, Vuong TD, Blair MW, Gaur PM, Nguyen HT and Varshney RK 2011 Trait mapping and molecular breeding; in *Biology and breeding of food legumes* (eds) A Pratap and J Kumar (CABI International) pp 296–313

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD and Varshney RK 2008 Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC Plant Biol.* **8** 55

Dubey A, Farmer A, Schlueter J, Cannon SB, Abernathy B, Tuteja R, Woodward J, Shah T, *et al.* 2011 Defining the transcriptome assembly and its use for genome dynamics and transcriptome profiling studies in pigeonpea (*Cajanus cajan* L.). *DNA Res.* **18** 153–164

Duranti M and Gius C 1997 Legume seeds: protein content and nutritional value. *Field Crop Res.* **53** 31–45

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

Gautami B, Ravi K, Lakshmi Narasu M, Hoisington DA and Varshney RK 2009 Novel set of groundnut SSRs for genetic diversity and interspecific transferability. *Int. J. Integr. Biol.* **7** 100–106

Gautami B, Pandey MK, Vadez V, Nigam SN, Ratnakumar P, Krishnamurthy L, Radhakrishnan T, Gowda MVC, *et al.* 2011 Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Mol. Breed.* doi 10.1007/s11032-011-9660-0

Glaszmann JC, Kilian B, Upadhyaya HD and Varshney RK 2010 Accessing genetic diversity for crop improvement. *Curr. Opin. Plant Biol.* **13** 1–7

Gnanesh BN, Bohra A, Sharma M, Byregowda M, Pandey S, Wesley V, Saxena RK, Saxena KB, *et al.* 2011 Genetic mapping and quantitative trait locus analysis of resistance to sterility mosaic disease in pigeonpea [*Cajanus cajan* (L.) *Millsp.*]. *Field Crop Res.* **123** 53–61

Grusak MA 2002 Enhancing mineral content in plant food products. *J. Am. College Nutr.* **21** 178S–183S

Gujaria N, Kumar A, Dauthal P, Dubey A, Hiremath P, Bhanu Prakash A, Farmer A, Bhide M, *et al.* 2011 Development and use of genic molecular markers (GMMs) for construction of a transcript map of chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **122** 1577–1589

Gupta PK and Varshney RK 2000 The development and use of microsatellite markers for genetics and plant breeding with emphasis on bread wheat. *Euphytica* **113** 163–185

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tuteja R, Kumar A, Bhanuprakash A, *et al.* 2011 Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnol. J.* **9** 922–931

Khedikar YP, Gowda MVC, Sarvamangala C, Patgar KV, Upadhyaya HD and Varshney RK 2010 A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **121** 71–84

Killian A, Huttner E, Wenzl P, Jaccoud D, Carling J, Caig V, Evers M, Heller-Uszynska K, *et al.* 2005 The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement; in, *In the wake of double helix: from the green revolution to the gene revolution* (Proceeding of the International Conference Bologna, Italy) pp 443–461

Kudapa H, Bharti AK, Cannon SB, Farmer AD, Mulaosmanovic B, Kramer R, Bohra A, Weeks NT, *et al.* 2012 A comprehensive transcriptome assembly of pigeonpea (*Cajanus cajan* L.) using Sanger and second-generation sequencing platforms. *Mol. Plant* doi: 10.1093/mp/ssr111

Kulwal PL, Thudi M and Varshney RK 2011 Genomics interventions in crop breeding for sustainable agriculture; in

Encyclopedia of sustainability science and technology (Springer) (in press)

Lewis G, Schrire B, Mackind B and Lock M 2005 *Legumes of the world* (Royal Botanic Gardens, Kew, England)

Mace ES, Varshney RK, Mahalakshmi V, Seetha K, Gafoor A, Leeladevi Y and Crouch JH 2007 *In silico* development of simple sequence repeat markers within the aeschynomenoid/dalbergoid and genistoid clades of the Leguminosae family and their transferability to *Arachis hypogaea*, groundnut. *Plant Sci.* **174** 51–60

Mallikarjuna N, Senapathy S, Jadhav DR, Saxena KB, Sharma HC, Upadhyaya HD, Rathore A and Varshney RK 2011 Progress in the utilization of *Cajanus platycarpus* (Benth.) Maesen in pigeonpea improvement. *Plant Breed.* **130** 507–514

Nayak SN, Zhu H, Varghese N, Choi HK, Datta S, Horres R, Jüngling R, Singh J, *et al.* 2010 Integration of novel SSR and gene-based marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theor. Appl. Genet.* 120 1415–1441

Neumann K, Kobiljski B, Dencic S, Varshney RK and Börner A 2011 Genome-wide association mapping: a case study in bread wheat (*Triticum aestivum* L.). *Mol. Breed.* 27 37–58

Pandey MK, Monyo E, Ozias-Akins P, Liang X, Guimarães P, Nigam SN, Upadhyaya HD, Janila P, *et al.* 2011 Advances in *Arachis* genomics for peanut improvement. *Biotechnol. Adv.* doi:10.1016/j.biotechadv.2011.11.001

Pandey MK, Gautami B, Jayakumar T, Sriswathi M, Upadhyaya HD, Gowda MVC, Radhakrishnan T, *et al.* 2012 Highly informative genic and genomic SSR markers to facilitate molecular breeding in cultivated groundnut (*Arachis hypogaea*). *Plant Breed.* doi:10.1111/j.1439-0523.2011.01911.x

Raju NL, Gnanesh BN, Pazhamala L, Jayashree B, Pande S, Hiremath PJ, Byregowda M, Singh NK, *et al.* 2010 The first set of EST resource for gene discovery and marker development in pigeonpea (*Cajanus cajan* L.). *BMC Plant Biol.* **10** 45

Ravi K, Vadez V, Isobe S, Mir RR, Guo Y, Nigam SN, Gowda MVC, Radhakrishnan T, *et al.* 2011 Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **122** 1119–1132

Roy SJ, Tucker EJ and Tester M 2011 Genetic analysis of abiotic stress tolerance in crops. *Curr. Opin. Plant Biol.* **14** 232–239

Sarvamangala C, Gowda MVC and Varshney RK 2011 Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). *Field Crop Res.* **122** 49–59

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

Saxena RK, Saxena KB, Kumar RV, Hoisington DA and Varshney RK 2010a SSR-based diversity in elite pigeonpea genotypes for developing mapping populations to map resistance to fusarium wilt and sterility mosaic disease. *Plant Breed.* **129** 135–141

Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, Nadaf HL, Gautami B, Sarvamangala C, *et al.* 2011 Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Mol. Breed.* doi: 10.1007/s11032-011-9661-z

Thudi M, Bohra A, Nayak SN, Varghese N, Shah TM, Penmetsa RV, Thirunavukkarasu N, Gudipati S, *et al.* 2011 Novel SSR markers from BAC-end sequences, DArT arrays and a comprehensive genetic map with 1,291 marker loci for chickpea (*Cicer arietinum* L.). *PLoS One* **6** e27275

Thudi M, Li Y, Jackson SA, May GD and Varshney RK 2012 Current state-of-art of sequencing technologies for plant genomics research. *Brief. Functional Genomics* doi:10.1093/bfgp/elr045

Upadhyaya HD and Ortiz R 2001 A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor. Appl. Genet.* **102** 1292-1298

Upadhyaya H D, Thudi M, Dronavalli N, Gujaria N, Singh S, Sharma S and Varshney R K 2011 Genomic tools and germplasm diversity for chickpea improvement. *Plant Genet. Res.* **9** 45–58

Vance CP, Graham PH and Allan DL 2000 Biological nitrogen fixation phosphorus: a critical future need; in *Nitrogen fixation: from molecules to crop productivity* (eds) FO Pedrosa, M Hungria, MG Yates and WE Newton (Dordrecht, The Netherlands: Kluwer Academic Publishers) pp 506–514

Varshney RK, Graner A and Sorrells ME 2005a Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* **10** 621–630

Varshney RK, Hoisington DA and Tyagi AK 2006 Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* **24** 490–499

Varshney RK, Hoisington DA, Upadhyaya HD, Gaur PM, Nigam SN, Saxena K, Vadez V, Sethy NK, *et al.* 2007 Molecular genetics and breeding of grain legume crops for the semiarid tropics; in *Genomics assisted crop improvement Vol II: Genomics applications in crops* (eds) RK Varshney and R Tuberosa (Dordrecht, The Netherlands: Springer) pp 207–242

Varshney RK, Bertioli DJ, Moretzsohn MC, Vadez V, Krishnamurty L, Aruna R, Nigam SN, Ravi K, He G, Knapp SJ and Hoisington DA 2009d The first SSR based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **118** 729–39

Varshney RK, Close TJ, Singh NK, Hoisington DA and Cook DR 2009b Orphan legume crops enter the genomics era. *Curr. Opin. Plant Biol.* **12** 202–210

Varshney RK, Hiremath PJ, Lekha PT, Kashiwagi J, Jayasree B, Deokar AA, Vadez V, Xiao Y, *et al.* 2009c A comprehensive resource of drought-and salinity-responsive ESTs for gene discovery and marker development in chickpea (*Cicer arietinum* L). *BMC Genomics* **10** 523

Varshney RK, Hoisington DA, Nayak SN and Graner A 2009a Molecular plant breeding: methodology and achievements; in *Methods in molecular biology: Plant genomics* (eds) D Somers, P Langridge and PJ Gustafson (Totowa, NJ: The Humana Press) pp 283–304

Varshney RK, Thudi M, May GD and Jackson SA 2010 Legume genomics and breeding. *Plant Breed Rev.* **33** 257–304

Varshney RK, Paulo MJ, Grando S, van Eeuwijk FA, Keizer LTP, Guo P, Ceccarelli S, Kilian A, Baum M and Graner A 2012a Genome

wide association analyses for drought tolerance related traits in barley (*Hordeum vulgare L.*). *Field Crop Res.***126** 171–180

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MT, Azam S, *et al.* 2012b Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* **30** 83–89

Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, *et al.* 2006 A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* **7** 206

Wilson RF, Stalker HT and Brummer EC 2004 *Legume crop genomics* (Champaign, Illinois: AOCS press) pp 1–362

Yang SY, Saxena RK, Kulwal P, Ash GJ, Dubey A, Harper JDI, Upadhyaya HD, Kilian A, *et al.* 2011 The first genetic map of pigeonpea (*Cajanus cajan*) based on Diversity Arrays Technology (DArT) markers. *J. Genet.* **90** 103–109