Development of cost-effective SNP assays for chickpea genome analysis and breeding

RK Varshney*, S Nayak, B Jayashree, K Eshwar, HD Upadhyaya and **DA Hoisington**
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India
*Corresponding author: r.k.varshney@cgiar.org

Although microsatellite or simple sequence repeat (SSR) markers have been the preferred markers for plant genetics and breeding, single nucleotide polymorphisms (SNPs) are the most common class that detect the smallest unit of genetic variation present in genomes (Rafalski 2002). Marker technologies exploiting the potential of SNPs provide the possibility of constructing genetic maps at 100-fold higher marker densities than with other types of DNA polymorphisms (Cho et al. 1999, Sachidanandam et al. 2001). Identification and mapping of SNPs has been initiated recently for crop species like rice (*Oryza sativa*) (Feltus et al. 2004), maize (*Zea mays*) (Tenaillon et al. 2001), wheat (*Triticum aestivum*) (Somers et al. 2003), barley (*Hordeum vulgare*) (Kota et al. 2001, Rostoks et al. 2005), ryegrass (*Lolium multiflorum*) (Miura et al. 2005) and rye (*Secale cereale*) (Varshney et al. 2007). Several platforms (eg, DHPLC, pyrosequencing, MALDI-TOF spectrophotometry) for detecting SNPs are available; however, these require expensive equipment or consumables and may result in considerable costs per data point. As an alternative, if a particular SNP involves a recognition site for a restriction enzyme, a cleaved amplified polymorphic sequence (CAPS) assay can be used for genotyping the SNP (Varshney et al. 2007). To date, in chickpea (*Cicer arietium*), there is one report available on the development of CAPS and dCAPS from Bacterial Artificial Chromosome (BAC)-end sequences (Rajesh et al. 2005). Our study utilizes the existing expressed sequence tag (EST) resource of *Cicer* species for mining the SNPs and subsequently converting the SNPs into CAPS assays.

A total of 1499 ESTs generated from 26 different *Cicer* species, available in the public domain at the time of analysis were used for *in silico* identification of SNPs using the bioinformatic tools developed at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) (http://hpc.icrisat.org/PBSWeb). Cluster analysis provided a total of 118 clusters, of which 11 clusters contained sequences from more than one *Cicer* species. Further, these clusters were assembled into 19 contigs and 184 putative SNPs were identified in 15 contigs. However, only 73 SNPs involved restriction enzyme sites for development of the CAPS assays as identified through the SNP2CAPs program. Primer pairs were designed for only 8 contigs (CL3a, CL3c, CL3d, CL3e, CL4a, CL10, CL20 and CL99) which had SNPs, resulting in putative recognition sites to commonly used restriction enzymes.

In order to validate *in silico* SNPs, a total of 12 genotypes representing 9 *Cicer* species – *C. pungens* (ICC 17138), *C. bijugum* (ICC 17122), *C. microphyllum* (ICC 17248), *C. judaicum* (ICC 17148), *C. cuneatum* (ICC 17162), *C. yamashitae* (ICC 17116), *C. pinnatifidum* (ICC 17152), *C. reticulatum* (ICC 17123 and PI 489777) and *C. arietinum* (ICC 8261, ICC 4958 and ICC 1882) – were genotyped with each primer pair. Since sometimes there were more than one restriction enzyme for assaying a primer amplified SNP, a total of 17 primer-restriction enzyme combinations using the common restriction enzymes – *Xmn*I, *Nla*III, *Acc*I, *Aci*I, *Ban*I, *Hpa*II, *Xba*I, *Taq*I, *Eco*RV, *Rsa*I, *Sal*I, *Bst*NI and *Hae*III – were tested.

Out of 17 primer-enzyme combinations, polymorphic restriction patterns showing fragments of varying length were observed in five combinations. The restriction pattern for the CL3e contig with *Aci*I enzyme has been shown in Figure 1, where the genotype ICC 17162 (*C. cuneatum*) yielded a larger fragment (400 bp) while for the remaining 11 genotypes, a smaller fragment (350 bp) was obtained. Table 1 presents the restriction fragment patterns for five CAPS markers across all 12 genotypes. Sequencing of some digested and undigested restriction fragments confirmed the SNPs at the sequence level as well (data not shown). However, the CAPS assay could not be optimized in the remaining 12 primer-enzyme combinations (70%) which is most likely due to not having been able to genotype all the species from which the SNPs were derived. We included all the *Cicer* species (for which seed germination and DNA isolation was possible) from the genebank at ICRISAT, although only three genotypes of cultivated chickpea (*C. arietinum*) – ICC 8261, ICC 4958 and ICC 1882 – were used in this study. None of the

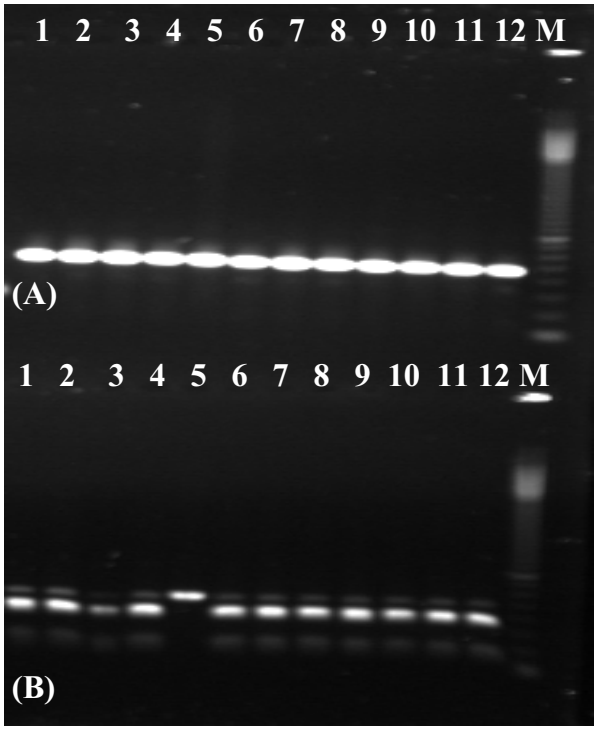

Figure 1. An example of CAPS assay for marker CL3e. Panels A and B represent the un-digested and digested (with *Aci*I) amplicons for 12 genotypes. The order of genotypes in gel is: 1 = ICC 17138 (*C. pungens*), 2 = ICC 17122 (*C. bijugum*), 3 = ICC 17248 (*C. microphyllum*), 4 = ICC 17148 (*C. judaicum*), 5 = ICC 17162 (*C. cuneatum*), 6 = ICC 17116 (*C. yamashitae*), 7 = ICC 17123 (*C. reticulatum*), 8 = ICC 17152 (*C. pinnatifidum*), 9 = ICC 8261 (*C. arietinum*), 10 = PI 489777 (*C. reticulatum*), 11 = ICC 4958 (*C. arietinum*), 12 = ICC 1882 (*C. arietinum*). M represents the DNA Standard, 100 bp ladder (Amersham Pharmacia Biotech, USA).

developed CAPS markers showed polymorphism between these genotypes thereby once again indicating the low levels of polymorphism within *C. arietinum.*

In summary, this study clearly demonstrates the utility of *Cicer* EST resources and the availability of bioinformatics analysis pipelines for the large-scale identification of SNPs on the HPC (High Performance Computer) at ICRISAT and the development of cost-effective CAPS assay for SNP genotyping. It is anticipated that the availability of large number of ESTs from more than one genotype of cultivated chickpea (*C. arietinum*) in the near future will make it possible to

Table 1. Restriction pattern in 12 *Cicer* genotypes with five CAPS markers.

Primer name	Restriction enzyme	Fragment size	ICC 17138	ICC 17122	ICC 17248	ICC 17148	ICC 17162	ICC 17116	ICC 17123	ICC 17152	ICC 8261	PI 489777	ICC 4958	ICC 1882
CL3e	*Aci*I	1=400 bp, 2=350 bp	2	2	2	2	1	2	2	2	2	2	2	2
CL3e	*Hae*III	1=300 bp, 2=250 bp	2	1	2	1	1	2	1	1	1	1	1	1
CL4a	*Eco*RV	1=225 bp, 2=200 bp	2	1	2	1	2	2	2	2	2	2	2	2
CL20	*Bst*NI	1=400 bp, 2=350 bp	2	–	2	1	–	1	1	–	–	1	1	–
CL99	*Xba*I	1=500 bp, 2=300 bp	1	1	1	1	2	1	1	1	1	1	1	1

develop larger number of SNPs in cultivated chickpea germplasm for genome analysis and breeding applications.

Acknowledgment. Financial support from the National Fund for Basic and Strategic Research (NFBSRA) of Indian Council of Agricultural Research (ICAR), New Delhi and Generation Challenge Programme (GCP) of Consultative Group on International Agricultural Research (CGIAR) is gratefully acknowledged.

References

Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkard E, Dewdney L, Reuber TL, Stammers M, Federspiel N, Theologis A, Yang WH, Hubbell E, Au M, Chung EY, Lashkari D, Lemieux B, Dean C, Lipshutz RJ, Ausubel FM, Davis RW and **Oefner PJ.** 1999. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. Nature Genetics 23:203–207.

Feltus FA, Wan J, Schulze SR, Estill JC, Jiang N and **Paterson AH.** 2004. An SNP resource for rice genetics and breeding based on subspecies *indica* and *japonica* genome alignments. Genome Research 14:1812–1819.

Kota R, Varshney RK, Thiel T, Dehmer KJ and **Graner A.** 2001. Generation and comparison of EST-derived SSR and SNP markers in barley (*Hordeum vulgare* L.). Hereditas 135:141–151.

Miura Y, Ding C, Ozaki R, Hirata M, Fujimori M, Takahashi W, Cai H and **Mizuno K.** 2005. Development of EST-derived CAPS and AFLP markers linked to a gene for resistance to ryegrass blast (*Pyricularia* sp.) in Italian ryegrass (*Lolium multiflorum* Lam.). Theoretical and Applied Genetics 111:811–818.

Rafalski JA. 2002. Application of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5:94–100.

Rajesh PN, McPhee K and **Muehlbauer FJ.** 2005. Detection of polymorphism using CAPS and dCAPS markers in two chickpea genotypes. International Chickpea and Pigeonpea Newsletter 12:4–6.

Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF and **Waugh R.** 2005. Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Molecular Genetics and Genomics 274:515–527.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE and **Cole CG.** 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933.

Somers DJ, Kirkpatrick R, Moniwa M and **Walsh A.** 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. Genome 46:431–437.

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF and **Gaut BS.** 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays* L.). Proceedings of the National Academy of Sciences, USA 98:9161–9166.

Varshney RK, Beier U, Khlestkina E, Kota R, Korzun V, Röder M, Graner A and **Börner A.** 2007. Single nucleotide polymorphisms in rye: discovery, frequency and applications for genome mapping and diversity studies. Theoretical and Applied Genetics 114:1105–1116.