Extending the repertoire of microsatellite markers for genetic linkage mapping and germplasm screening in chickpea

RK Varshney[1]*, R Horres[2], C Molina[3], S Nayak[1], R Jungmann[3], P Swamy[1], P Winter[2], B Jayashree[1], G Kahl[3] and **DA Hoisington[1]**

1. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India
2. GenXPro GmbH, Frankfurter Innovationszentrum Biotechnologie (FIZ), Altenhöferallee 3, D-60438 Frankfurt am Main, Germany
3. University of Frankfurt, Max von Laue Str. 9, D-60439 Frankfurt am Main, Germany

*Corresponding author: r.k.varshney@cgiar.org

Citation : Varshney RK, Horres R, Molina C, Nayak S, Jungmann R, Swamy P, Winter P, Jayashree B, Kahl G and Hoisington DA, (2007) Extending the repertoire of microsatellite markers for genetic linkage mapping and germplasm screening in chickpea. Journal of SAT Agricultural Resarch 5(1).

Molecular markers and genetic linkage maps are the prerequisites for undertaking molecular breeding activities. However, the progress towards development of a reasonable number of molecular markers has been very slow in cultivated species of chickpea (*Cicer arietinum*). One of the main reasons for this may be attributed to the low level of genetic diversity present in the cultivated gene pools of these species, at least with the detection tools that are currently available. Among various molecular markers currently available, microsatellite or SSR (simple sequence repeat) markers are often chosen as the preferred markers for a variety of applications in breeding because of their multi-allelic nature, co-dominant inheritance, relative abundance and extensive genome coverage (Gupta and Varshney 2000). As a result, several hundred of SSR markers have been developed in chickpea (Hüttel et al. 1999, Winter et al. 1999, Sethy et al. 2003, 2006, Lichtenzveig et al. 2005, Choudhary et al. 2006). Despite the availability of hundreds of SSR markers, the density of the intraspecific genetic maps is still very low. With an objective to increase the number of polymorphic SSR markers in chickpea, a genomic library was constructed and SSRs derived from this approach were characterized in this study.

A genomic DNA library from the chickpea genotype ICC 4958 was constructed after digesting total DNA of ICC 4958 with *MBO*I/*Sau* and *Taq*I at University of Frankfurt, Germany. Selected fragments of 800–1200 bp were purified and ligated to Promega pGEM 3Z(f) vector (Promega, USA). Plasmids were recovered and electroporated into *Escherichia coli* Sure strain (DH10B, Stratagene) at 1.8 kV/cm using a BIORAD Gene pulser and cloned in DH10B (*E. coli*). After plating about 400,000 clones at a density of 20,000 colonies per plate on 24 cm × 24 cm petri dishes (Nunc, Germany), masterplates [384-well micro-titre plates (MTPs)] were generated. Replicas were taken from these master plates and spotted in duplicate on (positively charged) PVDF-Macroarrays at RZPD GmbH, Berlin, Germany (http://www.rzpd.de). Subsequently, the hybridization of macroarrays with the 3'end-labeled synthetic repeats $(GA)_{10}$ and $(TAA)_{10}$ and screening of positive clones were performed at GenXPro GmbH. Clones with positive signals were grown on LB (Luria Bertani) agar plates with ampicillin overnight at 37°C. Aliquots of these colonies were used for colony PCR (polymerase chain reaction). Subsequently, a total of 359 clones with strong to moderate signals for the two SSRs were selected.

The bacterial colonies were grown overnight at 37°C in 5 ml of Luria Broth medium with 100 µg ml^{-1} ampicillin in an orbital shaker (Thermo Electron Corporation, USA) at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India. The plasmid DNA from individual clones was isolated using standard alkaline lysis method (Sambrook and Russell 2001). After checking the quality of the plasmid DNA on 0.8% agarose gel, the clones were sequenced at ICRISAT and Macrogen (South Korea) using the BigDye Terminator cycle sequencing kit on an ABI3700/ABI3730XL

Table 1. Summary on SSR search in the microsatellite enriched genomic library.

Total number of sequences examined	457
Total size of examined sequences (bp)	286,718
Total number of identified SSRs	643
Number of SSR containing sequences	299
Number of sequences containing more than one SSR	165
Number of SSRs present in compound form	221

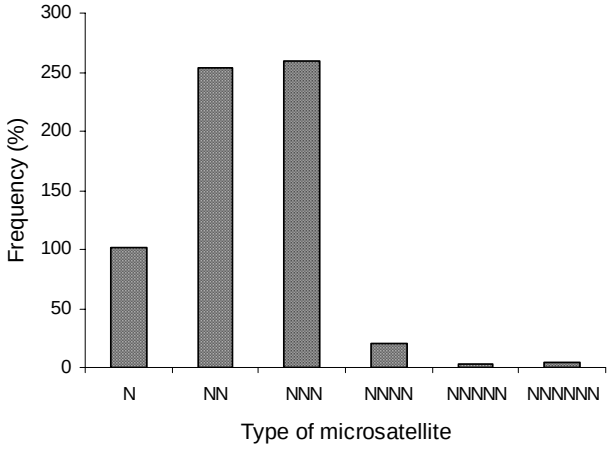

Figure 1. Frequency of different SSR classes.

(Applied Biosystems, USA). A total of 288 clones were sequenced at Macrogen in both directions using standard T7 promoter and SP6 primers and 576 sequences were generated. Nineteen clones were sequenced at ICRISAT in only one direction by using M13-forward sequencing primer. In total, 595 forward and/or reverse sequences were generated. Subsequently, all 595 raw sequences were trimmed using Sequencher program to remove the residues of the vector sequences following standard parameters. The cleaned sequences were checked for the quality of run, by inspecting the chromatograms of the sequence data. In 23 cases the chromatogram showed bad sequencing quality; hence those 23 sequence data were removed from the list. Further, the sequence residues less than 100 bp and more than 700 bp were removed in order to ensure high quality sequence data. Sequences obtained using the T7 promoter and SP6 primers were subjected to Contig Assembly Program (CAP3) to obtain the contig sequence (http://bioweb.pasteur.fr/seqanal/interfaces/cap3.html). Out of a total of 572 cleaned sequences (from forward as well as reverse), only 115 contigs could be generated, while 342 sequences remained singletons.

Ultimately, a total of 457 sequences representing 115 contigs and 342 singletons were screened for the presence of SSR using a Perl5 script called *MISA* (*MI*cro*SA*tellite), available at http://pgrc.ipk-gatersleben.de/misa/. The *MISA* search tool is capable of identifying both perfect and compound SSRs. Mononucleotide (N) SSR with minimum 10 repeat units and di- (NN), tri- (NNN), tetra- (NNNN), penta- (NNNNN) and hexa- (NNNNNN) nucleotide with a minimum of four repeat units were considered. The compound microsatellites were considered with a minimum distance of 100 bp. The sequence complementarity was considered while classifying the microsatellites. For instance, the repeat motifs AG, GC, TC, CT were grouped in the same class. Following the above mentioned criteria, a total of 643 SSRs were identified in 299 sequences out of 457 sequences representing 28.67 kb sequence data (Table 1). A total of 165 sequences contained more than one SSR while 221 sequences were found to have compound SSRs.

In terms of the frequency of different SSRs identified, the di- and tri-nucleotide SSRs were the most abundant (39% and 40%, respectively). Mononucleotide SSRs occurred in about 16% cases and tetranucleotide SSRs in about 3%. Other types of SSRs had <1% representation (Fig. 1). In majority (467; 73%) of the 643 SSR motifs, the microsatellites were of less than 10 repeat units followed by those having repeat units between 10 and 20

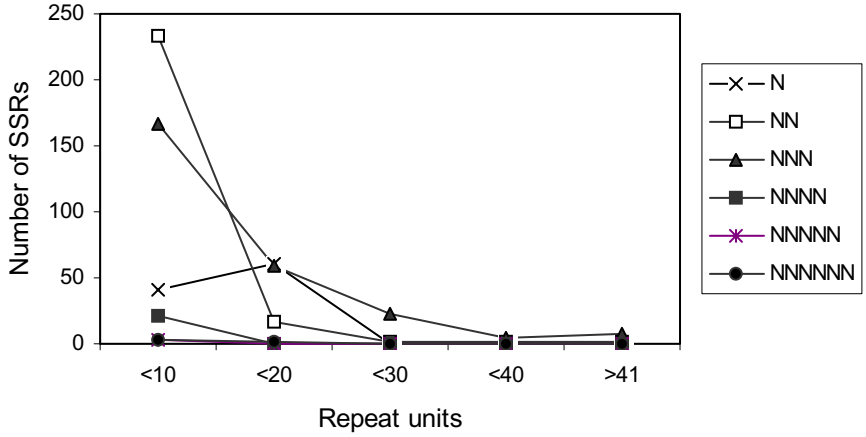

Figure 2. Number of SSRs with different number of repeat units.

(137 cases; 21%). However, only very few microsatellites were observed in other classes of SSR, eg, 20-30, 30-40, etc. The distribution of different lengths of SSRs is shown in Figure 2.

In order to develop SSR markers for chickpea genome analysis, all 299 sequences containing SSRs (including 165 sequences where more than one SSR was present) were used for primer design using the PRIMER3 program in a batch file. As a result, a total of 311 primer pairs were designed. For some sequences, primer pairs were designed for more than one SSR. Although this increases the redundancy, it ensures not to miss any polymorphic SSR in a given sequence.

At present, these SSR markers are being screened on a panel of 12 parental genotypes representing parental genotypes of six intraspecific (*C. arietinum* × *C. arietinum*) mapping populations – ICC 506EB × Vijay, ICC 3137 × IG 72953, ICC 3137 × IG 72933, ICC 283 × ICC 8261, ICC 4958 × ICC 1882, ICCV 2 × JG 62 – and one interspecific reference mapping population *C. arietinum* ICC 4958 × *C. reticulatum* PI 489777. Markers showing polymorphism in above mentioned combinations will be used to genotype the corresponding mapping populations to prepare linkage maps and trait mapping at ICRISAT. Another interspecific mapping population *C. arietinum* ICC 6098 × *C. reticulatum* ILWC 36 will be genotyped at Frankfurt. The developed resource of SSR markers is available upon request and has been shared with several collaborators including Washington State University, Pullman, USA (Fred Muehlbauer), The Australian Centre for Necrotrophic Fungal Pathogens, Murdoch University, Perth, Australia (Richard Oliver), the National Chemical Laboratory, Pune, India (Vidya Gupta) and the National Research Centre on Plant Biotechnology, New Delhi, India (Ramamurthy Srinivasan). In summary, this study increases the existing SSR repertoire in chickpea, which will help to enhance the coverage of linkage maps especially in intraspecific crosses where marker polymorphism is found to be very less.

Acknowledgment. We gratefully acknowledge the financial support received from Generation Challenge Programme through Tropical Legume I Project and Agri-Science Park of ICRISAT under the umbrella of Israel-ICRISAT collaborative grant.

References

Choudhary S, Sethy NK, Shokeen B and **Bhatia S.** 2006. Development of sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). Molecular Ecology Notes 6:3–95.

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

Hüttel B, Winter P, Weising K, Choumane W, Weigand F and **Kahl G.** 1999. Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). Genome 42:210–217.

Lichtenzveig J, Scheuring C, Dodge J, Abbo S and **Zhang HB.** 2005. Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea (*Cicer arietinum* (L.)). Theoretical and Applied Genetics 110:492–510.

Sambrook J and **Russell DW.** 2001. Molecular cloning. A laboratory manual. Edn 3. Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press. 540 pp.

Sethy NK, Choudhary S, Shokeen B and **Bhatia S.** 2006. Identification of microsatellite markers from *Cicer reticulatum*: molecular variation and phylogenetic analysis. Theoretical and Applied Genetics 112:347– 357.

Sethy NK, Shokeen B and **Bhatia S.** 2003. Isolation and characterization of sequence-tagged microsatellite sites markers in chickpea (*Cicer arietinum* L.). Molecular Ecology Notes 3:428–430.

Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ and **Kahl G.** 1999. Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*Cicer arietinum* L.) genome. Molecular and General Genetics 262:90–101.