Genomic variation studies in *Glycine max* and *Glycine soja* using SSR markers

Soybean [*Glycine max* (L.) Merr.] is an important source of edible oil and nutritious food worldwide. In India, it is the most important oilseed crop occupying an area of about 10.84 m ha with an average production and productivity of 14.66 mt and 1353 kg/ha respectively, during 2013–14 (http://www.agricoop.nic.in). However, yield of soybean in India is lower than the global average (2.5 t/ha). Low genetic diversity among Indian soybean germplasm, devoid of genetic resources in gene pool-2 (GP-2), various biotic and abiotic stresses, use of limited germplasm lines in breeding programmes, little or no use of wild-type germplasm in breeding, are some of the factors for lower yield in India. Genetic improvement, agronomic management and effective plant protection approaches are keys to increased soybean yield. However, yield is a complex, quantitative trait and is governed by quantitative trait loci or QTL. Therefore, for introgression of useful QTL from wild species, crosses were made between *G. max* and *Glycine soja*, and advanced segregating generation (F_6) has been developed. The specific objective of this work was to study the genetic relations and level of polymorphism between *G. max* and *G. soja*, and identifying polymorphic simple sequence repeats (SSR) markers. The genomic information and polymorphic markers identified in this study would be useful in analysing and mapping QTL for yield and other important traits.

DS9712, an Indian soybean variety was taken as representative of *G. max*, while a wild-type accession (DC2008-1) was taken as representative of *G. soja*. Genomic DNA was extracted from the young leaves of both the genotypes separately following CTAB (cetyl trimethyl ammonium bromide) procedure[1]. Quality and quantity of the DNA so extracted was estimated through gel electrophoresis and spectrophotometer analysis respectively. Purified DNA was subjected to PCR amplification in 20 µl reaction mixture containing 5 µl DNA (20 ng/µl), 2 µl PCR 10× buffer (with $MgCl_2$), 2 µl dNTPs (100 mM), 2 µl each of forward and reverse SSR primers (20 ng/µl), 0.3 µl *Taq* DNA polymerase (3U/µl) and 6.7 µl double-distilled water in a thermocycler. The DNA was denatured at 94°C for 4 min followed by 45 cycles each consisting of denaturation at 94°C for 1 min, primer annealing at 49–55°C for 1 min, primer elongation at 72°C for 1 min and final elongation at 72°C for 10 min. Amplified products so obtained were resolved in 3% metaphore agarose gel stained with ethidium bromide and analysed in gel imaging system.

For genomic study, 262 SSR markers (@13 markers/chromosome) were selected at random from across the soybean genome. The sequences of the SSR markers were obtained from genomic resources available on-line (www.soybase.org) and synthesized locally[2]. Among the markers used, 172 were found to be polymorphic between the two genotypes (Figure 1), indicating the level of genomic variations between the two species to be 65.67%. In previous studies, the level of genetic variations within *G. max* was reported[3] to be 43.38–48.38%. Thus though *G. soja* and *G. max* are cross-compatible, they maintain enormous genetic variations. The genomic variations were also reflected in the phenotypic trait variations. Plants of the cultivated species are erect with sturdy stem and determinate growth, whereas the *G. soja* plants are climbers with indeterminate growth (Figure 2). *G. soja* flowers are very small and purple in colour, while *G. max* flowers are relatively big and white or purple in colour. The seeds of *G. soja* are black, whereas those of *G. max* may be black, yellow, tawny, etc. The pods of *G. soja* are small with tiny seeds that shutter easily upon maturity. However, *G. soja* is resistant to a number of biotic and abiotic stresses, including YMV

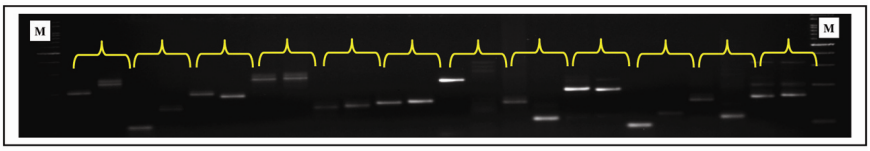

Figure 1. Amplification patterns of 12 SSR markers in soybean genome. In each pair of lanes, left and right side lanes represent *Glycine max* and *Glycine soja* genotypes respectively. The markers in the gel from left to right are Satt418 (ATT_{27}), Satt063 (ATT_{20}), Satt601 (ATT_{13}), Satt663 (ATT_{28}), Satt070 (ATT_{24}), Satt656 (ATT_{10}), Sat_238 (AT_{28}), Satt460 (ATT_{26}), Satt685 (ATT_{14}), Satt618 (ATT_{14}), Satt308 (ATT_{21}), and Satt687 (ATT_9). Figures in the parentheses (e.g. ATT) indicate repeat motif and numerals (e.g. 24, 26, etc.) represent the number of repeats. M, 100-bp ladder.

Figure 2. Morphological variation in plant type of (***a***) *G. max* cv. DS 9712 and (***b***) *G. soja*.

Table 1. Chromosome-wise distribution of SSR markers used and their level of polymorphism obtained between the two species

Chromosome no.*	SSR markers used (no.)	SSR polymorphic (no.)	SSR monomorphic (no.)	Polymorphism (%)	Remarks#
1 (D1a)	17	8	9	47.06	Low
2 (D1b)	18	10	8	55.56	Medium
3 (N)	15	9	6	60.00	Medium
4 (C1)	16	10	6	62.50	Medium
5 (A1)	18	11	7	61.11	Medium
6 (C2)	10	6	4	60.00	Medium
7 (M)	11	6	5	54.55	Medium
8 (A2)	12	8	4	66.67	Medium
9 (K)	14	12	2	85.71	High
10 (O)	10	9	1	90.00	High
11 (B1)	14	7	7	50.00	Medium
12 (H)	13	11	2	84.62	High
13 (F2)	12	7	5	58.33	Medium
14 (B2)	10	7	3	70.00	Medium
15 (E2)	10	6	4	60.00	Medium
16 (J)	15	10	5	66.67	Medium
17 (D2)	16	9	7	56.25	Medium
18 (G)	10	8	2	80.00	High
19 (L)	11	9	2	81.82	High
20 (I)	10	9	1	90.00	High

*Figure in parentheses indicates linkage group; #High: > 70%, medium: 50–70%; low: <50%.

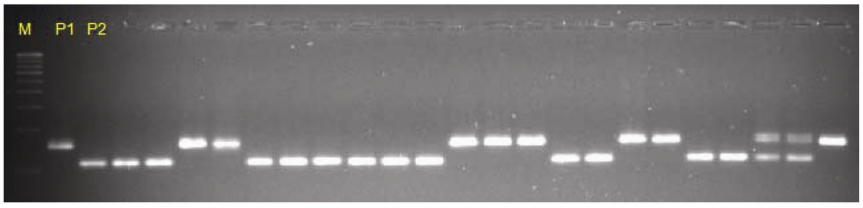

Figure 3. Pattern of bands in the inter-specific hybrid-derived population of *G. max* and *G. soja*. The P1 (*G. max*) and P2 (*G. soja*) alleles segregated in F_6 generation plants. Lane 1, Marker: lane 2, P1; lane 3, P2; lanes 4–22, Segregating plants.

disease which is serious in northern India. The useful traits of *G. soja* can be introgressed into the cultivated species through hybridization. Successful introgression of QTL for useful traits from *G. soja* into *G. max* has already been reported[4]. The polymorphic markers identified in this study were used to analyse the molecular genotype of the segregating plants of the inter-specific crosses. As expected, the parental alleles were in 1 : 1 ratio (Figure 3). However, heterozygosity was also found in some loci indicating the need for more selfing generations to make it a true recombinant inbred line (RIL). The RIL produced from this study would be used for mapping and analysing the QTLs for important traits, including yield.

The distribution of polymorphic SSR loci was not uniform across the genome. Some chromosomes had more polymorphic loci than others (Table 1). Highest level of polymorphism (90%) was observed in chromosome nos 10 and 19, while the least (47.06%) was observed in chromosome no. 1. It indicated that a number of genomic regions are still conserved between the two species; however, distributions of such regions are random across the genome. More the conserved regions, less is the level of polymorphism and vice versa. Chromosomes with lower level of polymorphism need to be tested with more number of markers per unit length to find polymorphic loci.

The power of SSR markers to find genomic variations was found to vary with the type of motif and the number of repeats in a motif. The trinucleotide motifs could detect more variations (54.06% polymorphism) than the di (37.79%) or polynucleotide motifs (8.13%). Therefore, while designing new primers, it would be prudent to pick those which would amplify the trinucleotide repeats over others. With more than three nucleotides in a repeat motif coupled with higher number of repeat copies, the level of polymorphism goes down. Further, it was noticed that the level of polymorphism detected between the two species also varied with the number of times a particular motif is repeated. Out of the 90 monomorphic markers, 66 had repeat motif with less than 20 repeats. On the other hand, out of the 172 polymorphic markers, 91 had motifs with 20–30 or more repeats. Thus, it appears that SSRs with 20–30 repeats, i.e. say ATT_{20}, ATT_{25}, ATT_{30}, etc. are more polymorphic than others. It can be correlated with

the basic principle of polymorphism in microsatellite markers. Since the polymorphism detected through SSR markers is length polymorphism (simple sequence length polymorphism, SSLP), more the variation in the length amplified, more will be the polymorphism. Segment with more number of repeats carries better chances of producing polymorphism than the shorter ones. Moreover, larger variations are easily detected in gel electrophoresis. Thus, SSRs with 20–30 repeats are found to be more useful in detecting genetic variations between the two species.

The study reveals that enormous genetic variations still exist between *G. max* and *G. soja*. It also highlights some useful properties of SSR markers that would help in studying genetic variability in soybean genotypes. Further, the polymorphic markers identified in this study would be highly useful in mapping of QTL for yield and other important agronomic traits in soybean.

1. Saghai-Maroof, M. A., Soliman, K. M., Jergensen, R. A. and Allard, R. W., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 8014–8018.
2. Hwang, T. *et al.*, *DNA Res.*, 2009, **16**, 213–225.
3. Kumar, V. *et al.*, *Indian J. Genet.*, 2011, **71**, 372–376.
4. Concibido, V. C. *et al.*, *Theor. Appl. Genet.*, 2003, **106**, 575–582.

Received 23 January 2015; revised accepted 3 September 2015

YASHPAL[1]
D. R. RATHOD[1]
JYOTI DEVI[2]
ANIL KUMAR[1]
KEYA MUKHERJEE[1]
DEEPIKA CHERUKU[1]
SUBHASH CHANDRA[1]
S. K. LAL[1]
AKSHAY TALUKDAR[1,*]

[1]*Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India*
[2]*ICAR-Indian Vegetable Research Institute, Varanasi 221 005, India*
**For correspondence.*
e-mail: atalukdar@iari.res.in

Alkaline lamprophyre (camptonite) from Bayyaram area, NE margin of the Eastern Dharwar Craton, southern India

The widespread occurrence of lamprophyres is known, since more than a few decades, from the various parts of Cudappah Igneous Provice (CIP)/Prakasham Alkaline Province (PAP) of the Eastern Dharwar Craton (EDC), southern India (Figure 1 *a*). On inspecting the distribution of lamprophyres in three cratons, i.e. EDC, Aravalli–Bundelkhand and Bastar–Bhandar Craton[1], it is evident that the EDC alone hosts the maximum number and variety of lamprophyres[2]. The present correspondence reports a lamprophyre dyke near the Bayyaram area (80°09′:17°35′) at the northeastern margin of the EDC (Figure 1 *a*). It also addresses the petrology, geochemistry and significance of this occurrence.

The study area mainly consists of granitoids of Peninsular Gneissic Complex (PGC-II) of the EDC. Regionally, the area is bounded by two Proterozoic sedimentary basins, i.e. Pakhal basin to the east and Cuddapah basin to the south (Figure 1 *a*). The lamprophyre of the study area has been intruded within granitoids of the EDC. The dyke shows NW–SE trend having 15–20 m length and 1 m width approximately (Figure 1 *b*). Megascopically, the dyke is mesocratic to melanocratic, fine-grained and shows porphyritic texture. Phenocrysts are uniformly distributed in the fine-grained

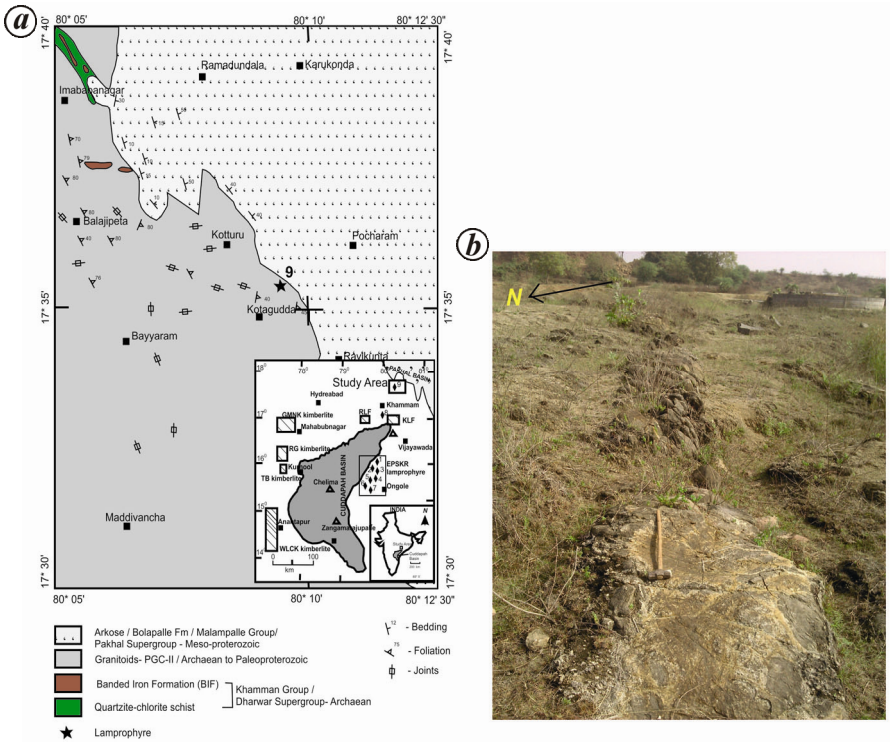

Figure 1. ***a***, Geological map of the Bayyaram area. (Inset) Map showing location of lamprophyre–lamproite–kimberlite fields in the Eastern Dharwar Craton (EDC) of South India (GMNK, Gulbarga, Maddur, Narayanpet Kotakonda Kimberlites; RG, Raichur Gawal Kimberlites; TB, Tungabhadra kimberlites; WLCL, Wajrakarur, Lattavaram, Chigicherla and Kalyanadurg kimberlites; RLF, Ramadugu lamproite field; KLF, Krishna lamproite field; ESKPPKR, Elchuru (1), Settupalle (2), Kommalapadu (3), Purimetla (4), Pusupugullu (5), Kellampalle (6), Ravipadu (7) lamprophyres in Prakasam district and Polayapalle (8), Bayyaram (9) Lamprophyres in Khamman district). ***b***, Field photograph showing outcrop pattern of lamprophyre dyke.