Indian J. Genet., 79(4) 693-703 (2019)
DOI: 10.31742/IJGPB.79.4.7

Deploying inter-specific recombinant inbred lines to map QTLs for yield-related traits in soybean

Yashpal, D. R. Rathod, Subhash Chandra, Anil Kumar, Raju Ratan Yadav and Akshay Talukdar*

Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi 110 012

(Received: July 2019; Revised: October 2019; Accepted: November 2019)

Abstract

Quantitative trait loci (QTL) mapping and analyses were conducted for yield and six yield-related traits in soybean using 184 inter-specific recombinant inbred lines (RILs) derived from a cross involving wild type (*Glycine soja* Sieb. & Zucc.) accession DC2008-1 and cultivated (*Glycine max* L. Merr.) variety DS9712. A Linkage map of 1639.55 cM length was constructed with 167 SSR markers (65.65% polymorphism) with an average marker interval of 9.82 cM. Using three years phenotypic data 34 QTLs were mapped for 7 traits using Inclusive Composite Interval Mapping approach. The number of QTLs mapped for a trait varied from year to year, however, QTLs for days-to-50% flowering (*qDFF5*), 100-seed weight (*qHSW9-1*, *qHSW9-2* and *qHSW19*) and yield (*qYLD17*) were mapped consistently over the three years of testing. Identified QTLs were validated through single marker analysis in 92 germplasm lines. The study demonstrated the potential of wild type soybean to harness QTLs for yield-related traits. The identified QTLs could be utilized for genetic improvement of soybean through molecular breeding.

Key words: QTLs, RIL, soybean, SSR markers, yield-related traits

Introduction

Soybean is an important oilseed crop of the world. It is the premier oilseed crop in India occupying more than 12 million hectare under cultivation; however, its average productivity (~1.0t/ha) is far below the world average (~2.5t/ha). The escalating demand for soy-based food products due to an increased awareness about its health benefits and multifarious applications is asking for enhanced production and productivity of soybean. Soybean in India which was introduced during 1963 from the USA (Agarwal et al. 2013) maintains a narrow base gene pool (Delannay et al. 1983; Gizlice et al. 1994). Further, the wild-type germplasm which constitute an underutilized hidden reservoir of useful genes (Wang et al. 1992; Lark et al. 1995; Maughan et al. 1996; Mian et al. 1996; Orf et al. 1999a, b) has been least used in soybean genetic studies and breeding programmes in India (Yashpal et al. 2015).

Mapping of quantitative trait loci (QTL) constitutes an important activity in genetic improvement of soybean. The availability of QTL linked marker targeting desirable traits offers a faster and more accurate approach to breeding, through genotype-based selection (Concibido et al. 2003; Guang-yu et al. 2011). Previously, the use of molecular markers have allowed identification and mapping of many soybean QTLs for economically important traits such as days-to-50% flowering (Orf et al. 1999b); number of pods per plant, 100-seed weight (Vieira et al. 2006); days-to-maturity, plant height (Yuan et al. 2002; Eskandari et al. 2013) etc. These studies primarily used mortal mapping populations such as F_2 and BC with limited population size and non-repetition of the study over different environments (years, locations or both). Contrarily, in this study, an inter-specific recombinant inbred line (RIL) mapping population has been used in a replicated trial over 3 years. The RILs offers the advantages of 'true-breeding' homozygous lines (Collard et al. 2005) and repetition over the years of experimentation represents different environments (Hackett et al. 2010). The parental genotypes used for developing the RILs were highly contrasting for several characters and hence the RILs contained mosaic genome and offered better polymorphism between loci. Yield and related traits being quantitative in nature with low heritability warrants use of large mapping population for accuracy of analysis and

*Corresponding author's e-mail: akshay.talukdar1@gmail.com
Published by the Indian Society of Genetics & Plant Breeding, A-Block, F2, First Floor, NASC Complex, IARI P.O., Pusa Campus, New Delhi 110 012; Online management by www.isgpb.org; indianjournals.com

precision of mapping. Quantitative traits are highly affected by the growing environments hence repetition of the trial over different environments is crucial to realise the consensual marker-trait associations. Therefore, the present study used 184 inter-specific RILs which were phenotyped over three years and genotyped with SSR markers to identify QTL for yield and its attributing traits.

Materials and methods

Plant materials and phenotyping

A mapping population comprising of 184 inter-specific RILs ($F_{2:8}$) developed by single-seed descent method was used in the study. The RILs were developed from the cross, DC2008-1/DS9712. The genotype DC2008-1 was an accession of wild-type soybean (*G. soja*) while DS9712 was a popular variety of cultivated soybean (*G. max*). Both the genotypes varied for a number of traits *viz*., plant type (indeterminate, climber vs determinate, erect), flower colour (purple vs white), maturity duration (138 vs 117 days), seed size (small vs large), seed colour (black vs yellow), etc. The pods of DC2008-1 were small with tiny black seeds (100-seed weight of 0.54-0.56 g) that shattered easily upon maturity; while DS9712 possessed larger pods with yellow, bold seeds (100-seed weight of 7.9-8.4 g). Per plant yield of DS9712 was more (20.0-22.0 g) compared to DC2008-1 (0.67-0.68 g).

The RILs along with the parents were evaluated in the research farm of Indian Agricultural Research Institute at New Delhi (28°38'N, 77°10'E; 224 m) consecutively for three years i.e. 2013 (F_6), 2014 (F_7) and 2015 (F_8). Every year, the crop was raised in different plots following an augmented block design with 4 blocks and 6 checks (DS9712, DC2008-1, Bragg, VLS61, SL688 and PS1347). The parental genotypes and the RILs were grown in a plot of three-row of 1 m length. All the recommended agronomic management practices were followed to raise a healthy crop. The row-to-row distance was 45 cm and plant-to-plant distance was maintained at 10 cm. Data on per plant yield and six yield-related traits *viz*., days-to-50% flowering (DFF), days-to-maturity (DM), plant height (PH), number of primary branches per plant (PBPL), number of pods per plant (PDPL) and 100-seeds weight (HSW) were recorded from the parents and RILs. Some of the RILs inherited the trait of pod-shattering from the wild type parent (DC2008-1) which could have a masking effect on yield. Therefore, manual threshing was used to separate the seeds from pods once they attained physical maturity.

Molecular genotyping

In order to study the level of polymorphism between the parents, 262 SSR markers distributed evenly across the soybean genome were picked up from the consensus genetic map (Cregan et al. 1999). The sequences of the selected SSR markers were downloaded from the *SoyBase* website (www.soybase.org), and synthesized through local vendors.

Genomic DNA was extracted from the young leaves of the parents and RILs following CTAB (cetyl trimethyl ammonium bromide) procedure (Saghai-Maroof et al. 1984). Purified DNA was subjected to PCR amplification in 20 µl reaction mixture containing 5.0 µl template DNA (20 ng/µl), 2.0 µl *Taq* buffer A-with $MgCl_2$ (10×), 2.0 µl dNTPs (25 mM), 2.0 µl each forward and reverse SSR primers (30 ng/µl), 0.3 µl *Taq* DNA polymerase (3U/µl) and 6.7µl double distilled water. Amplification of the template DNA was performed in a thermocycler (Applied Biosystem) in following steps. Step I: denaturation at 94°C for 4 min.; Step II: denaturation at 94°C for 1 min.; primer annealing at 49-55°C for 1 min.; primer elongation at 72°C for 1 min. Step II was repeated for 35 cycles. Step III: Final elongation at 72°C for 10 min. Amplified products were resolved on 3% Metaphore™ gel stained with Ethidium Bromide in 1xTAE buffer at 80V for 2.5-3.0 hours and analyzed in a gel documentation system (AlphaImager-ProtienSimple).

Statistical analyses

The quantitative data were tested for normality using MS Excel. Pearson's correlation coefficients among the yield and six yield-related traits were obtained using a web-based statistical analysis service of Indian Agricultural Statistics Research Institute, New Delhi (Rathore et al. 2004). Mean, range and standard deviation (SD) of the data pertaining to yield and yield-related traits of the RILs and of the validation population were analyzed using SAS version 9.4. ANOVA was prepared as per the augmented block design of experiments. The treatment sum of square was partitioned into sum of square among RILs, among checks and between RILs and checks. The adjusted means of RILs and checks were obtained and used for further analyses. The critical difference was obtained at 5% level of significance for all the four different kinds of comparisons *viz*., between checks, between RILs and checks, between RILs of the same block and between RILs of different blocks as per

Federer (1956); Federer (1961); Parsad and Gupta (2000).

Linkage map construction and QTL mapping

Out of the 262 SSR markers used for parental polymorphism survey, 172 SSRs appeared to be polymorphic. These polymorphic SSRs were used for genotyping of 184 RILs along with parental lines. Genotypic data of the RILs were subjected to a goodness of fit test using χ^2 statistics for checking segregation distortion. Out of 172 polymorphic markers, 5 markers were rejected due to complex banding pattern and distorted segregation, and the remaining 167 markers were used to construct the linkage map. The map construction was carried out in the QTL IciMapping software v 4.1.0 (Meng et al. 2015) using the linkage mapping function. A logarithm of the odds (LOD) threshold of 3.0 was used for grouping. For the ordering of markers within each linkage group SERiation algorithm (Buetow and Chakravarti 1987) was used with the criteria SARF for rippling. The recombination frequencies were translated into genetic map distances using Kosambi's mapping function (Kosambi 1944).

The QTL analyses for each year of testing were performed separately for all the seven traits. The scanning interval was set to 1 cM, and the probability of stepwise regression for markers entering the model was set to 0.001. Adjusted means of the phenotypic data were used to map the QTL effects such as log-likelihood ratio (LOD) score, per cent phenotypic variation explained (PVE) and the additive effect of the QTL loci using the ICIM procedure (Wang 2009). A QTL was considered consistent if it was detected in at least two of the three years of testing. QTL nomenclature was followed as per McCouch et al. (1997).

Population validation

To validate the results of QTL analyses, a working mini-core collection comprising of 92 diverse genotypes was established based on the yield and yield-attributing traits for which RILs were phenotyped. This sub-set of germplasm lines consisted of released varieties, elite lines, exotic and indigenous genotypes including inter-specific derivatives of *G. max* and *G. soja*. The validation population was raised during 2014 and 2015 in a completely randomized block design (CRBD) with three replications and data were recorded for all the traits under study.

Results

Phenotypic performance and correlation studies

The results of the statistical analysis of the yield and yield-attributing traits are shown in Table 1. Significant variations were observed between the parental genotypes and among the RILs for yield and other yield-related traits (ANOVA not shown). Except PDPL, HSW and YLD, the wild type parent, DC2008-1, had higher values for DFF, DM, PH and PBPL than those of DS9712. Among the RILs, transgressive segregants were observed for all the traits

Table 1. Performance of yield and yield-attributing traits in parental genotypes and RILs over three years of testing

Traits	Parents						RILs					
	DC2008-1			DS9712			Year 2013		Year 2014		Year 2015	
	2013	2014	2015	2013	2014	2015	Range	Mean±SE	Range	Mean±SE	Range	Mean±SE
DFF	74.0	75.0	75.0	55.00	61.0	60.0	35-67	54.2±6.0	38-99	67.6±8.3	33-80	59.8±7.2
DM	137.0	138.0	138.0	111.00	119.0	121.0	76-113	98.6±6.2	94-127	111±6.9	85-132	114.8±8.7
PH	136.0	137.0	140.0	77.00	73.0	75.0	28-89	56.9±13.7	16-105	58±20.1	19-110	57.3±20.2
PBPL	8.60	8.70	8.90	2.80	2.8	2.9	2.4-13.6	6.8±1.9	0.6-12	5.9±2.4	0.4-14.3	5.5±2.6
PDPL	65.0	67.0	63.0	102.00	100.0	105.0	23-365	156.6±74.3	7.0-451	211±98	22-450	212±103
HSW	0.54	0.54	0.56	7.93	8.40	8.30	1.1-3.9	2.2±0.5	1.0-3.1	2.2±0.4	1.2-4.5	2.3±0.7
YLD	0.67	0.68	0.67	9.50	14.1	13.2	1.5-16.0	8.0±3.4	1.45-14.0	8.1±3.3	1.5-19.2	10.2±4.6

DFF: Days-to-50% flowering; DM: Days-to-maturity; PH: Plant height; PBPL: Primary branches per plant; PDPL: Number of pod per plant; HSW: 100-seeds weight; YLD: Yield per plant

except HSW. The data found to distribute normally for all the traits under study. Pearson's correlation coefficient indicated the existence of positive and negative correlations (P<0.01) among the traits studied (Table 2). Days-to-50% flowering (DFF) had a positive and significant correlation with all the traits except

Table 2. Correlations among yield and six related traits in the RILs

Traits	DFF	DM	PH	PBPL	PDPL	HSW
DM	0.498**					
PH	0.327**	0.306**				
PBPL	0.313**	0.137	0.433**			
PDPL	0.247**	0.074	0.440**	0.552**		
HSW	-0.627**	-0.073	-0.261**	-0.407**	-0.300**	
YLD	-0.092	0.144	0.253**	0.201**	0.395**	0.326**

**. Correlation is significant at the 0.01 level (2-tailed). DFF: Days-to-50% flowering; DM: Days-to-maturity; PH: Plant height; PBPL: Primary branches per plant; PDPL: Number of pods per plant; HSW: 100-seeds weight; YLD: Yield per plant

HSW and per plant yield (YLD). The yield had a positive correlation with all traits but DFF. Other traits also showed positive and significant correlations among themselves except HSW which showed a negative correlation with all the traits except YLD. No correlation was observed for DM vs PBPL, PDPL, HSW and YLD.

Parental polymorphism and linkage map construction

Out of the 262 random SSR markers used, 172 appeared to be polymorphic between the parental genotypes indicating a 65.65% level of polymorphism between the parental genotypes. Out of 172 polymorphic markers, 5 markers were rejected due to complex banding patterns and distorted segregations. The remaining 167 markers were deployed to develop a map of 1639.55 cM length with an average marker distance of 9.82 cM. Order of the markers on the map was the same as in the reference map (Cregan et al. 1999) with occasional variation in inter-marker map distance in a few cases. Further, barring one region on chromosome 17, all the chromosomes had an almost uniform distribution of the markers.

Mapping of QTLs

Using ICIM, a total of 34 QTLs were mapped over the three years of testing. The number of QTLs mapped varied from year to year. For example, 13 QTLs were mapped in 2013 while 9 and 12 QTLs were mapped during 2014 and 2015, respectively (Table 3). Number of QTLs detected for a particular trait varied from 1 to 5. In 2013, only one QTL was detected for DM and YLD while five QTLs were mapped for HSW. Similarly, one QTL each was mapped for DFF, DM and PH during 2014. The number of QTL mapped for HSW were thirteen across the three years of testing. These QTLs were detected only on 11 out of 20 soybean chromosomes (Fig. 1).

So far consistency of QTL is concerned, a limited number of QTL were found to be consistent over the years. For DFF, five QTLs were detected over the three years of testing; however, only one QTL (*qDFF5*) appeared consistently. It had a higher LOD value (above 3) and PVE ranged from 9.52-16.09% (Table 3). Similarly, for yield per plant, out of 3 QTLs, one (*qYLD17*) appeared to be consistent with PVE ranging from 10.59 to 13.02%. For HSW, four QTLs *viz*., *qHSW7*, *qHSW9-1*, *qHSQ9-2* and *qHSW19* appeared consistently over all the three years of testing. PVE by these QTLs ranged between 6.44 and 14.14%. For PH, three different QTLs were mapped on three different chromosomes; only one QTL (*qPH2*) was detected consistently during 2014 and 2015 with PVE ranging from 13.25 to 14.49%. Similar was the case with DM where only one QTL (*qDM5*) was found to be consistent in two out of the three years of testing. Two QTLs were mapped for PDPL and one for PBPL; however, none was consistent in its expression over the years (Table 3). In most of the consistent QTLs, the PVE appeared to vary with the year of testing. The QTL for days-to-maturity (*qDM17*) and yield (*qYLD17*) appeared to be coincident QTL (Table 3 and Fig. 1). Similarly, the QTL for DFF, *qDFF5* which appeared consistently over two years was mapped close to the PBPL QTL i.e. *qPBPL5* (Table 3 and Fig. 1).

The RILs and the set of SSR markers used for the QTL mapping analyses were the same over three years of testing. Therefore, a comparison of the chromosomal regions carrying common QTLs was made to identify the most consistent ones. There were five regions on four chromosomes that carried QTLs for HSW. On chromosome 7, the QTL, *qHSW7* was mapped at a position 110.85cM as defined by the marker pair Sat_121 and Satt618. Chromosome 9 carried two QTLs, *qHSW9-1* and *qHSW9-2*. The QTL *qHSW9-1* was mapped at 15.85 cM position flanked by markers Sat_087 and Sat_119, and the *qHSW9-2* was mapped at 52.85 cM position and was flanked by

Table 3. QTL for yield and yield-attributing traits identified during 2013-2015

Trait	Year	QTL	Chr. No. (LG)	Marker interval	Map position(cM)	LOD	PVE (%)	Add.
Days-to-50% flowering	2013	*qDFF2-1*	2 (D1b)	Satt266-Satt282	75.74	6.77	13.22	2.19
		qDFF5	5 (A1)	Satt174-Satt200	91.56	3.45	9.52	1.86
	2014	***qDFF5***	5 (A1)	Satt174-Satt200	92.56	6.63	16.09	3.38
	2015	*qDFF2-2*	2 (D1b)	Satt701-Satt634	45.74	3.81	8.70	2.12
		qDFF7	7M)	Satt618-Sat-_276	126.86	4.01	8.63	2.11
		qDFF10	10 (O)	Sat_282-Satt331	93.25	3.42	6.30	1.80
Days-to-maturity	2013	*qDM17*	17(D2)	Satt464-Sat_001	90.15	3.37	8.90	-1.93
	2014	*qDM5*	5(A1)	Satt593-Satt591	29.56	3.68	9.53	2.21
	2015	*qDM2*	2(D1b)	Satt296-Satt266	52.74	3.56	7.69	2.44
		qDM5	5(A1)	Satt591-Sat_356	31.56	3.38	7.40	2.38
Plant height (cm)	2013	*qPH5*	5(A1)	Satt591-Sat_356	40.56	3.04	7.50	3.73
		qPH12	12(H)	Satt676-Sat_175	82.59	3.35	7.81	-3.80
	2014	***qPH2***	2(D1b)	AI856415-Satt296	51.74	4.49	13.25	6.54
	2015	***qPH2***	2(D1b)	AI856415-Satt296	51.74	4.62	14.49	7.82
Number of primary branches/plant	2015	*qPBPL5*	5(A1)	Satt236-Satt511	93.56	4.00	13.29	0.98
Number of pods/plant	2013	*qPDPL3*	3(N)	Sat_379-Satt009	8.32	3.90	12.46	-28.90
		qPDPL12	12(H)	Satt192-Satt676	56.59	3.48	17.83	34.41
100-seed weight (g)	2013	***qHSW7***	7(M)	Sat_121-Satt618	110.85	4.12	7.10	-0.52
		qHSW9-1	9(K)	Sat_087-Sat_119	15.85	5.35	10.47	-0.63
		qHSW9-2	9(K)	Sat_116-Satt499	52.85	6.42	11.59	-0.66
		qHSW11	11(B1)	Satt251-Satt197	36.95	3.02	5.18	-0.44
		qHSW19	19(L)	Satt006-Sat_245	114.31	5.58	10.03	-0.62
	2014	*qHSW7*	7(M)	Satt618-Sat_276	111.85	4.44	7.83	-0.54
		qHSW9-1	9(K)	Sat_087-Sat_119	15.85	4.95	9.65	-0.60
		qHSW9-2	9(K)	Sat_116-Satt499	52.85	7.69	14.14	-0.73
		qHSW19	19(L)	Satt006-Sat_245	114.3	5.18	9.34	-0.59
	2015	***qHSW7***	7(M)	Sat_121-Satt618	110.85	4.89	8.19	-0.56
		qHSW9-1	9(K)	Sat_087-Sat_119	14.85	3.25	6.44	-0.50
		qHSW9-2	9(K)	Sat_116-Satt499	52.85	5.82	10.13	-0.63
		qHSW19	19(L)	Satt006-Sat_245	114.31	4.52	7.79	-0.55
Yield/plant (g)	2013	***qYLD17***	17(D2)	Satt464-Sat_001	90.15	5.47	13.02	-1.43
	2014	*qYLD16*	16(J)	Sat529-Sat_255	41.94	4.55	9.06	1.47
		qYLD17	17(D2)	Satt464-Sat_001	90.15	4.94	10.59	-1.60
	2015	*qYLD17*	17(D2)	Sat_365-Satt464	89.16	3.47	12.63	-1.67

QTLs in boldface are consistent ones across years. PVE is the percentages of phenotypic variance explained by the detected QTL

the markers Sat_116 and Satt499. Another HSW QTL, *qHSW19* was mapped on chromosome 19 at 114.31 cM position flanked by the markers Satt006 and Sat_245. Similarly, *qDFF5* also appeared consistently in the same map interval between 91.56 cM and 92.56 cM on chromosome 5, flanked by the markers Satt174 and Satt200. The QTL mapped for DM, *qDM5*, was located at the map position of 29.56 cM on chromosome 5 flanked by the markers Satt593 and Satt591 during 2014, which however, was located at a position 31.56 cM between the markers, Satt591 and Sat_356 during 2015. It may be possible that this QTL may be a single one or two different QTLs located around the marker Satt591 (Fig. 1). For other traits, QTL mapped varied from one year to another reflecting their interactions with the environment.

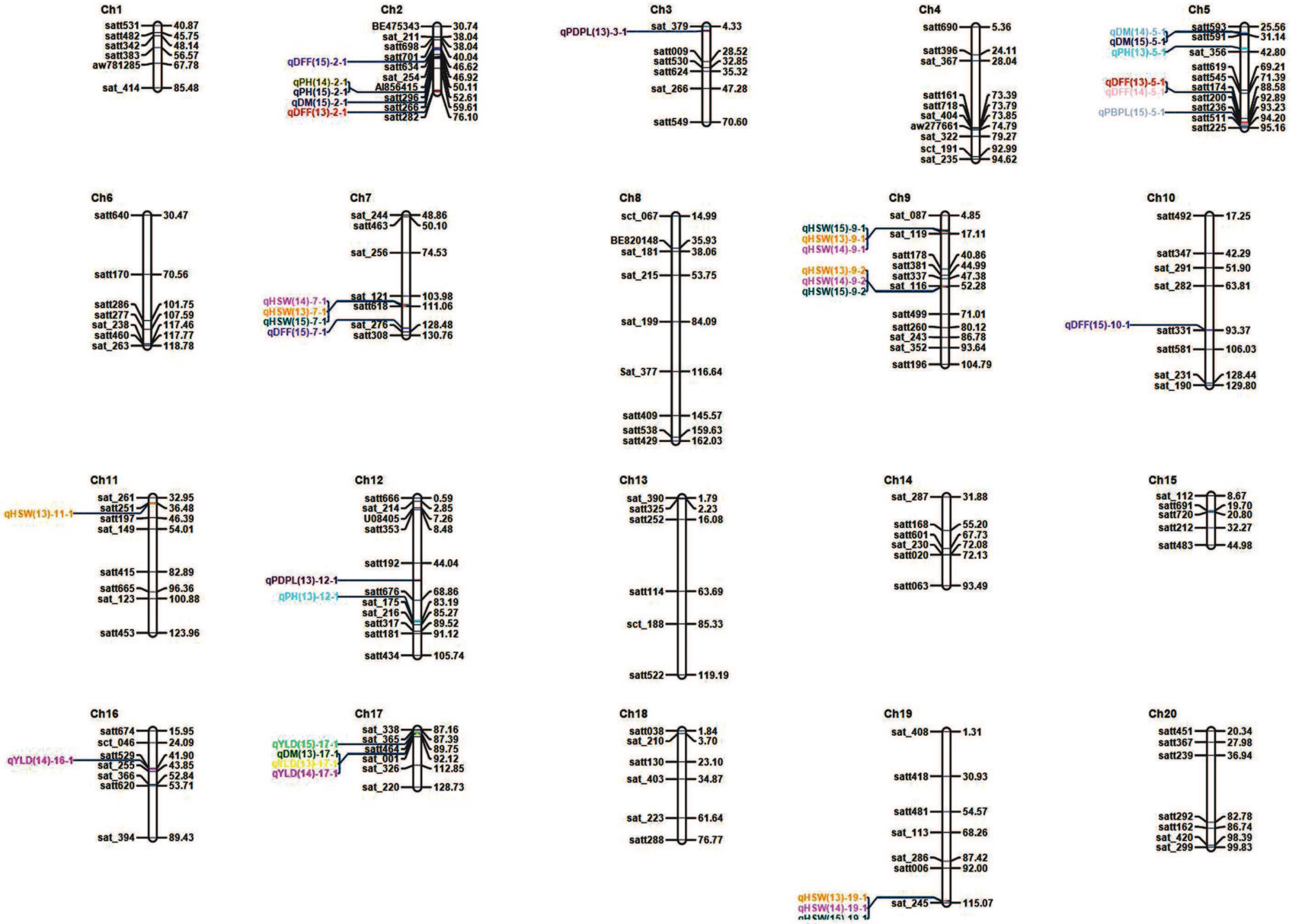

Fig. 1. Map position of various QTLs detected for seven traits in the study

Validation of the QTLs

Validation of the QTLs identified in the study was performed in a set of 92 germplasm lines through single marker analysis. Two QTLs were validated for DFF with a PVE of 18.69 and 22.16% followed by one QTL each for DM (PVE of 17.03%), PH (PVE of 12.23%) and YLD (PVE of 22.30%) (Table 4). Essentially, the

Table 4. QTL validated in germplasm lines through single marker analysis

Chr. No.	QTL	Map position (cM)	Marker	LOD	PVE (%)	ADD
2	*qDFF2*	76.103	satt282	5.00	22.16	3.24
5	*qDFF5*	92.887	satt200	4.13	18.69	2.66
5	*qDM5*	31.144	satt591	3.73	17.03	2.21
5	*qPH5*	31.144	satt591	2.60	12.23	5.35
17	*qYLD17*	89.752	satt464	5.04	22.30	-2.1

DFF: Days-to-50% flowering; DM: Days-to-maturity; PH: Plant height; PBPL: Primary branches per plant; PDPL: Number of pods per plant; HSW: 100-seeds weight; YLD: Yield per plant

QTLs that got validated in the germplasm lines were consistent QTLs. The effects of the QTL in the validation population were marginally higher than the RILs, with the similar contributing alleles. The QTL, *qDFF5* had additive value 1.86 in the RIL population which changed to 2.66 in the validation population. Similarly, *qYLD17* had additive value –1.43 (2013) to –1.67 (2015) in the RIL which got changed to –2.1 in the validation population.

Discussion

Plant breeders have successfully utilized the wild relatives for improvement of simply inherited traits like disease resistance; however, exploiting the wild species for productivity-related traits has emerged as an effective approach to enrich the genetic diversity of elite cultivars (Gaikwad et al. 2014). In this study, two sexually compatible parental genotypes *viz*., *G. soja* (DC2008-1) and *G. max* (DS9712) were used to develop an inter-specific mapping population consisting of 184 RILs (Kumar et al. 2019). This population was enormously diverse for several traits (Yashpal et al. 2015) including yield and components (Table 1) and included transgressive segregants for most of the traits. Zhang et al. (2004) and Li et al. (2008) also noted occurrence of transgressive segregants in the segregating population. Most of the traits exhibited significantly positive correlations among themselves, except for HSW which was found negatively associated with DFF, PH, PBPL and PDPL. Thus, most of the yield associated traits had apparent positive trend towards yield, indicating their usefulness in direct selection for yield in the RIL population. Li et al. (2008) have reported a negative correlation of seed weight with flowering duration and number of branches.

While genotyped with 262 random genome-wide SSRs, the parents showed higher level of polymorphism (65.65%) signifying the potential of exploiting inter-genomic variations. This was also apparent by the mosaic status of the RIL genomes encompassing both *G. max* and *G. soja* genomic components as evidenced by the expected 1:1 ratio of the parental allele segregation (Fig. 2) with the rare

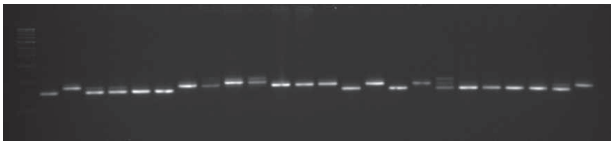

Fig. 2. Pattern of bands for Satt591 in the inter-specific RIL population. Lane 1: Marker (100-bp); Lane 2: P1 (DC2008-1; 160-bp); Lane 3: P2 (DS9712; 180-bp); Lanes 4-22: RILs. The parental alleles have shown 1:1 segregation ratio

occurrence of heterozygous alleles. Further, it indicated the suitability of the population for QTL mapping studies. The 1639.55 cM long linkage map with 167 polymorphic SSR markers deployed at an average interval of 9.82 cM was comparable with previously reported soybean linkage maps. Li (2006) constructed a genetic map that contained 121 markers and had a total length of 1506 cM with an average marker interval of 12.5 cM. While mapping QTL for domestication related traits, Liu et al. (2007) constructed a linkage map with 282 markers that spanned a length of 2383 cM with an average marker distance of 8.5 cM. Similarly, Li et al. (2008) constructed a genetic map with a total length of 1073.9 cM with an average distance between the adjacent marker loci of 7.9 cM. All such genetic maps including the one constructed in the present study are shorter than the integrated soybean linkage map (2524 cM) (Cregan et al. 1999). Such variations in map length are caused by a number of factors including the number of markers used, segregation pattern of the markers, missing values, accuracy of the linkage analysis, and coverage of the markers over the genome etc. Linkage map developed using large mapping population and

co-dominant markers are generally considered precise (Kumawat et al. 2012). The density of markers on the genetic maps found to vary from one to another. The density of the markers also depends on the level of homozygosity and recombination frequencies between the parental lines used to develop the mapping population (Castiglioni et al. 1999).

The influence of the environment (years) on the expression of QTL was visibly reflected in variations in the number of QTLs mapped and their PVE. Under significant genotype × environment interactions, it is possible that QTL for a particular trait may be mapped in different map positions across the environments. Contextual effects of environment and genotypes on dissection of QTLs has been reported previously in rice (Wang et al. 2014), maize (Boer et al. 2007; Li et al. 2018), barley (Zhao et al. 2012) and mustard (Ramchiary et al. 2007; Rout et al. 2015). This is pertinent in the present study because relatively fewer QTLs were detected during 2014 (9 nos.) as compared to 2013 (13 nos.) and 2015 (12 nos.) signifying the influence of environment on trait expression. This was also evident from the significant variation among the number of QTLs detected for a trait with years. Traits with low heritability tends to show such variation in QTL expression across environments. Among the QTLs mapped, a few such as *qDFF5*, *qHSW9-1*, *qHDW9-2*, *qHSW19* and *qYLD17* were detected consistently over the three years indicating their stability of expression, higher heritability, low QTL × environment interaction and reliability for application in the future breeding programs. Consistent QTLs are useful to a wide range of environment, whereas, environment-specific QTLs can be used within specific target environment (Li et al. 2015).

Wide variation in the number of QTLs mapped per trait and their corresponding phenotypic variance indicated that the genetic diversity of the population used could divulge different genetic loci governing yield and yield-related traits. It appeared that desirable alleles for various quantitative traits were present on either of the parental genotypes. Among the consistent QTLs, the trait enhancing alleles were contributed by the cultivated type parent. Breeding for earliness and higher productivity is imperative in soybean improvement. Earliness is determined by days to flowering and maturity. There were five QTLs mapped for days-to-50% flowering on chromosomes 2, 5, 7 and 10 with positive contribution from the wild-type parent DC2008-1. This implied that the *G. soja* alleles tend to delay the duration while, *G. max* alleles are beneficial for selection for earliness. In earlier studies, QTLs for DFF have been previously reported on Chr. 2 (Komatsu et al. 2007; Khan et al. 2008; Li et al. 2008), Chr.7 (Mansur et al. 1993; Mansur et al. 1996; Orf et al. 1999b) and Chr.11 (Reinprecht et al. 2006; Gai et al. 2007). The QTL on chromosome 5, *qDFF5* has been consistent across the two years. No such QTL has yet been reported yet on Chr.5. Therefore, *qDFF5* can be considered to be a novel QTL. Similarly, three QTLs for DM were mapped on chromosomes 2, 5 and 17, among which *qDM17* was considered to be a novel QTL as it has not yet been reported elsewhere. However, *qDM17* was mapped only during 2013 season. Among the three QTLs identified for plant height *viz*., *qPH2*, *qPH5*, and *qPH12*, the *qPH2* was the most robust one explaining 14.49% of the total variation. These QTLs are in conformity with previously reported QTLs for plant height. QTLs reported for branching are comparatively fewer in number and very few studies have been conducted on this trait (Sayama et al. 2010). As per the S*oyBase* database (www.soybase.org accessed on 10 July 2019), 21 QTLs have so far been reported for branching. In this study, one QTLs *viz*., *qPBPL5* was identified for this trait during 2015. The *G. soja* parent, DC2008-1 was the contributor for branch increasing allele for the QTL, *qPBPL5*. Similarly, two QTLs *viz*., *qPDPL3* and *qPDPL12* were identified for the PDPL during 2013 period. *qPDPL12* had positive allelic contribution from the *G. soja* parent. Zhang et al. (2010) had earlier reported mapping of a QTL for this trait on chromosome 3; therefore, *qPDPL12* could be a new QTL. Four of the seven consistent QTLs identified in this study were mapped for hundred seed weight (HSW). All of them had positive alleles contributed by the *G. max* parent, DS9712. For YLD also, two QTLs were identified with moderate to large phenotypic effect, of which one, *qYLD17* was highly consistent across years. The *qYLD17* has its yield contributing allele coming from the cultivated type soybean. Use of these consistent QTLs, provide opportunities to enhance seed yield through their deployment. However, despite its agronomically inferior appearance, *G. soja* accession DC2008-1 was shown to harbour trait enhancing QTL alleles for yield and its attributing traits with varying effect, which can be used in the improvement of elite soybean cultivars. The potential of exploiting un-adapted and wild germplasm for the trait improvement of elite cultivars has been reported previously by Tanksley et al. (1996) in tomato, Concibido et al. (2003) in soybean, and Gaikwad et al. (2014) in rice. The

novel QTLs can be subjected for fine mapping and map-based cloning, whereas, the QTLs that are mapped to chromosomal regions consistent with previous studies are useful for marker-assisted introgression into elite cultivars (Marri et al. 2005).

QTLs being highly sensitive to the environment and genetic backgrounds, its validation usually refer to the verification of its effectiveness under different genetic backgrounds (Langridge et al. 2001). Therefore, identified QTLs are tested in unrelated genotypes or mapping population and their efficacies and usefulness are determined. The validation test done on 92 germplasm lines validated two QTLs for DFF and one QTL each for DM, PH and YLD with their significant phenotypic variation for the corresponding traits in the test population. Furthermore, the PVE by these QTLs were higher in the test population. Palomeque et al. (2010) reported that the estimates of QTL is dependent on the strength of association with the traits of interest, allelic diversity at the QTL loci and the effect of the genetic background on expression of the QTL. Moreover, the mapping approach adopted needed to be taken into consideration while validation because PVE under single marker analysis usually appear higher than that of composite interval mapping. This study while demonstrating the potential of *G. soja* as a source of additional variability for improving traits of complex inheritance like yield, affirms the perspective that the wild relatives constitute an important but untapped reservoir of productivity-related alleles in soybean. Consistency of expression over the years, validation in unrelated germplasm and co-localization with previously reported QTLs confirms the robustness of several QTLs detected in the present study and indicates their future applicability in the soybean improvement programs through molecular breeding approach. Additionally, soybean genome sequence information (Schmutz et al. 2010) would facilitate the identification of genes underlying these QTLs. Utilization of various RILs generated in this study for crossing with elite cultivars would augment the variability of the existing gene pool and strengthen the Indian soybean varietal development programme.

Authors' contribution

Conceptualization of research (AT); Designing of the experiments (AT, YP, AK); Contribution of experimental materials (AT); Execution of field/lab experiments and data collection (YP, AK, RRY, SC); Analysis of data and interpretation (YP, DRR, AT); Preparation of manuscript (YP, AT).

Declaration

The authors declare no conflict of interest.

Acknowledgments

Yashpal is grateful to the Indian Council of Agricultural Research (ICAR), New Delhi for Senior Research Fellowship during the doctoral study. Authors are grateful to the National Agricultural Science Fund, ICAR, New Delhi for funding the program.

References

Agarwal D. K., Billore S. D., Sharma A. N., Dupare B. U. and Srivastava S. K. 2013. Soybean: Introduction, improvement, and utilization in India-problems and prospects. Agric. Res., **2:** 293-300.

Boer M. P., Wright D., Feng L., Podlich D. W., Luo L., Cooper M. and van Eeuwijk F. A. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics, **177:** 1801-1813.

Buetow K. H. and Chakravarti A. 1987. Multipoint gene mapping using seriation. I. General methods. Am. J. Hum. Genet., **41:** 180-188.

Castiglioni P., Ajmone-Marsan P., Van Wijk R. and Motto M. 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor. Appl. Genet., **99:** 425-431.

Collard B. C., Jahufer M. Z., Brouwer J. B. and Pang E. C. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, **142:** 169-196.

Concibido V., La Vallee B., Mclaird P., Pineda N., Meyer J., Hummel L., Yang J., Wu K. and Delannay X. 2003. Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. Theor. Appl. Genet., **106:** 575-82.

Cregan P. B., Jarvik T., Bush A. L., Shoemaker R. C., Lark K. G., Kahler A. L., Kaya N., Vantaoi T. T , Lohnes D. G., Chung J. and Specht J. E. 1999. An integrated genetic linkage map of soybean. Crop Sci., **39:** 1464-1490.

Delannay X., Rodgers D. M. and Palmer R. G. 1983. Relative genetic contributions among ancestral lines to North American soybean cultivars. Crop Sci., **23:** 944-949.

Eskandari M., Cober E. and Rajcan I. 2013. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. Theor. Appl. Genet., **126:** 1677-1687.

Federer W. T. 1956. Augmented designs. Hawaiian Planter

Record, **55:** 191-208.

Federer W. T. 1961. Augmented designs with one way elimination of heterogeneity. Biometrics, **17:** 447-473.

Gai J., Wang Y., Wu X. and Chen S. 2007. A comparative study on segregation analysis and QTL mapping of quantitative traits in plants with a case in soybean. Front. Agric. China, **1:** 1-7.

Gaikwad K. B., Singh N., Bhatia D., Kaur R., Bains N. S., Bharaj T. S. and Singh K. 2014. Yield-enhancing heterotic QTL transferred from wild species to cultivated rice *Oryza sativa* L. PLoS One, **9:** e96939.

Gizlice Z., Carter T. E. Jr and Burton J. W. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Sci., **34:** 1143-1151.

Guang-yu G., Rui S., Meng H., Yong-xin G., Da-wei X., Hong-wei J., Chun-yan L., Guo-hua H. and Qing-shan C. 2011. Quantitative trait locus (QTL) analysis of pod related traits in different environments in soybean. Afr. J. Biotechnol., **10:** 11848-11854.

Hackett C. A., Russell J., Jorgensen L., Gordon S. L. and Brennan R. M. 2010. Multi-environment QTL mapping in blackcurrant (*Ribes nigrum* L.) using mixed models. Theor. Appl. Genet., **121:** 1483-1488.

Khan N., Githiri S., Benitez E., Abe J., Kawasaki S., Hayashi T. and Takahashi R. 2008. QTL analysis of cleistogamy in soybean. Theor. Appl. Genet., **117:** 479-487.

Komatsu K., Okuda S., Takahashi M., Matsunaga R. and Nakazawa Y. 2007. Quantitative trait loci mapping of pubescence density and flowering time of insect resistant soybean. Genet. Mol. Biol., **30:** 635-639.

Kosambi D. D. 1944. The estimation of map distances from recombination values. Ann Eugen., **12:** 172-175.

Kumar A., Chandra S., Talukdar Akshay, Yadav R. R., Saini M., Poonia S. and Lal S. K. 2019. Genetic studies on seed coat permeability and viability in RILs derived from an inter-specific crosses in soybean. Indian J. Genet., **79**(1): 48-55.

Kumawat G., Raje R. S., Bhutani S., Pal J. P., Mithra A. S. V. R. C., Gaikwad K., Sharma T. R. and Singh N. K. 2012. Molecular mapping of QTLs for plant type and earliness traits in pigeonpea (*Cajanus cajan* L. Millsp.). BMC Genet., **13:** 84.

Langridge P., Lagudah E. S., Holton T. A., Appels R., Sharp P. J. and Chalmers K. J. 2001. Trends in genetic and genome analyses in wheat: a review. Aus. J. Agric. Res., **52:** 1043-1077.

Lark K. G., Chase K., Adler F., Mansur L. M. and Orf J. H. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. Proc. Natl. Acad Sci. U. S. A., **92:** 4656-4660.

Li D. 2006. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. Unpublished Ph. D. Thesis. University of Kentucky, Lexington, USA.

Li P., Zhang Y., Yin S., Zhu P., Pan T., Xu Y., Wang J., Hao D., Fang H., Xu C. and Yang Z. 2018. QTL-by-environment interaction in the response of maize root and shoot traits to different water regimes. Front Plant Sci., **9:** 229.

Li S., Wang J. and Zhang L. 2015. Inclusive composite interval mapping of QTL by environment interactions in biparental populations. PLoS One, **10:** e0132414.

Li W., Zheng D., Van K. and Lee S. 2008. QTL mapping for major agronomic traits across two years in soybean (*Glycine max* L. Merr.). J. Crop Sci. Biotechnol., **11:** 171-190.

Liu B., Fujita T., Yan Z. H., Sakamoto S., Xu D. and Abe J. 2007. QTL mapping of domestication related traits in soybean (*Glycine max*). Ann. Bot., **100:** 1027-1038.

Mansur L. M., Lark K. G., Kross H. and Oliveira A. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor. Appl. Genet., **86:** 907-913.

Mansur L. M., Orf J. H., Chase K., Jarvik T., Cregan P. B. and Lark K. G. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci., **36:** 1327-1336.

Marri P. R., Sarla N., Reddy L. V. and Siddiq E. A. 2005. Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. BMC genet., **6:** 33.

Maughan P. J., Saghai Maroof M. A. and Buss G. R. 1996. Molecularmarker analysis of seed-weight: genomic locations gene action and evidence for orthologous evolution among three legume species. Theor. Appl. Genet., **93:** 574-579.

McCouch S. R., Cho Y. G., Yano M., Paul E. and Blinstrub M. 1997. Report on QTL nomenclature. Rice Genet. Newsl., **14:** 11-13.

Meng L., Li H., Zhang L. and Wang J. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. The Crop J., **3:** 269-283.

Mian M. A. R., Bailey M. A., Tamulonis J. P., Shipe E. R., Carter T. E. Jr, Parrott W. A., Ashely D. A., Hussey R. S. and Boerma H. R. 1996. Molecular markers associated with seed weight in two soybean populations. Theor. Appl. Genet., **93:** 1011-1016.

Orf J. H., Chase K., Adler F. R., Mansur L. M. and Lark K. G. 1999a. Genetics of soybean agronomic traits. II. Interactions between yield quantitative trait loci in soybean. Crop Sci., **39:** 1652-1656.

Orf J. H., Chase K., Jarvik T., Mansur L. M., Cregan P. B., Adler F. R. and Lark K. G. 1999b. Genetics of soybean agronomic traits. I. Comparison of three related recombinant inbred populations. Crop Sci., **39:** 1642-1651.

Palomeque L., Liu L. J., Li W., Hedges B. R., Cober E. R., Smid M. P., Lukens L. and Rajcan I. 2010. Validation of mega-environment universal and specific QTL associated with seed yield and agronomic traits in soybeans. Theor. Appl. Genet., **120:** 997-1003.

Parsad R. and Gupta V. K. 2000. A Note on augmented designs. Indian J. Pl. Genet. Resour., **13:** 53-58.

Ramchiary N., Bisht N. C., Gupta V., Mukhopadhyay A., Arumugam N., Sodhi Y. S., Pental D. and Pradhan A. K. 2007. QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed Brassica juncea: importance of recurrent selection backcross scheme for the identification of 'true' QTL. Theor. Appl. Genet., **116:** 77-85.

Rathore A., Parsad R. and Gupta V. K. 2004. Computer aided construction and analysis of augmented designs. J Indian Soc. Agric. Stat., **57**(SV): 320-344 http://www.iasri.res.in/spadweb/default.aspx. Accessed 30 November 2015.

Reinprecht Y., Poysa V., Yu K., Rajcan I., Ablett G. and Pauls K. 2006. Seed and agronomic QTL in low linolenic acid, lipoxygenasefree soybean (*Glycine max* (L.) Merrill) germplasm. Genome, **49:** 1510-1527.

Rout K., Sharma M., Gupta V., Mukhopadhyay A., Sodhi Y. S., Pental D. and Pradhan A. K. 2015. Deciphering allelic variations for seed glucosinolate traits in oilseed mustard (*Brassica juncea*) using two bi-parental mapping populations. Theor. Appl. Genet., **128:** 657-666.

Saghai-Maroof M. A., Soliman K. M., Jergensen R. A. and Allard R. W. 1984. Ribosomal DNA spaces-length polymorphism in barley Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. U. S. A., **81:** 8014-8018.

Sayama T., Hwang T., Yamazaki H., Yamaguchi N., Komatsu K., Takahashi M., Suzuki C., Miyoshi T., Tanaka Y., Xia Z., Tsubokura Y., Watanabe S., Harada K., Funatsuki H. and Ishimoto M. 2010. Mapping and comparison of quantitative trait loci for soybean branching phenotype in two locations. Breed. Sci., **60:** 380-389.

Schmutz J., Cannon S. B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D. L., Song Q., Thelen J. J., Cheng J. and Xu D. 2010. Genome sequence of the palaeopolyploid soybean. Nature, **463:** 178-183.

Tanksley S. D., Grandillo S., Fulton T., Zamir D., Eshed T., Petiard V., Lopez J. and Beck-Bunn T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor. Appl. Genet., **92:** 213-224.

Vieira A., de-Oliveira A., Soares T., Schuster I., Piovesan N., Martinez C., de Barros E. and Moreira M. 2006. Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. Braz. J. Plant Physiol., **18:** 281-290.

Wang J. 2009. Inclusive composite interval mapping of quantitative trait genes. Acta Agronomica Sinica, **35:** 239-245.

Wang X., Pang Y., Zhang J., Zhang Q., Tao Y., Feng B., Zheng T., Xu J. and Li Z. 2014. Genetic background effects on QTL and QTL × environment interaction for yield and its component traits as revealed by reciprocal introgression lines in rice. The Crop J., **2:** 345-57.

Wang Z. Y., Second G. and Tanksley S. D. 1992. Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. Theor. Appl. Genet., **83:** 565-581.

Yashpal, Rathod D. R., Devi J., Kumar A., Mukherjee K., Cheruku D., Chandra S., Lal S. K. and Talukdar A. 2015. Genomic variation studies in *Glycine max* and *Glycine soja* using SSR markers. Curr. Sci., **109:** 1929-1931.

Yuan J., Njiti V. N., Meksem K., Iqbal M. J., Triwitayakorn K., Kassem M. A., Davis G. T., Schmidt M. E. and Lightfoot D. A. 2002. Quantitative trait loci in two soybean recombinant inbred line populations segregating for yield and disease resistance. Crop Sci., **42:** 271-277.

Zhang D., Cheng H., Wang H., Hengyou Z., Liu C. and Yu D. 2010. Identification of genomic regions determining flower and pod numbers development in soybean (*Glycine max* L.). J. Genet. Genomics, **37:** 545-556.

Zhang W. K., Wang Y. J., Luo G. Z., Zhang J. S., He C. Y., Wu X. L., Gai J. Y. and Chen S. Y. 2004. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. Theor. Appl. Genet., **108:** 1131-1139.

Zhao F. and Xu S. 2012. Genotype by environment interaction of quantitative traits: a case study in barley. G3: Genes, Genomes, Genetics, **2:** 779-788.