Indian J. Genet., 80(4) 475-479 (2020)
DOI: 10.31742/IJGPB.80.4.15

Short Communication

Identifying QTL for seed weight in a cross between vegetable and grain type soybeans

Gaurav Khosla*, B. S. Gill, Asmita Sirari, Pankaj Sharma and Inderpreet Dhaliwal

Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana 141 004

(Received: Auguest 2020; Revised: October 2020; Accepted: November 2020)

Abstract

The variation for 100 seed weight in F_2 population derived from a cross, AGS456 (an exotic vegetable type from Taiwan)/SL958 of soybean followed a normal curve with a range of 8.00-27.22g indicating quantitative nature of genetic control for seed size. Parental lines were screened with 207 SSR markers to identify polymorphism and 90 primer pairs detected polymorphism between the parents. These ninety markers were used for detecting polymorphism between two extreme bulks for seed weight. Out of these, 18 primer pairs were polymorphic for the bulks and were used for bulk segregant analysis in 200 F_2 plants. One major QTL for seed weight was identified on LG M with Sat_244 and Satt175 as flanking markers, explaining 19.0 per cent phenotypic variation. Two minor QTLs were also identified on LG D1b, one in interval Satt041-Sat_069 with an estimated phenotypic variation (R^2) of 6.0 per cent and the other in interval Sat_069-Sat_0183 estimating 7.0 per cent phenotypic variance. The markers flanking TLs may help in marker-assisted selection (MAS) for improvement of seed weight in soybean after fine mapping and validation.

Key words: Soybean, 100-seed weight, BSA, QTL, SSR markers

Soybean [*Glycine max* (L.) Merr.] is an important legume grown for food, feed and industrial uses. Seed size in soybean is a major target of breeding, not only as a component of seed yield (Liang et al. 2005) but also as a major trait while developing food grade soybean (Cui and Xuan 2007). Soybeans can be classified in two main groups: grain type and food type. Grain soybeans produce medium-sized seeds, represented by the 100-seed weight averaging 15g whereas vegetable soybeans are large seeded, ideally > 30g per 100 seeds (Zhang et al. 2010). The seeds of vegetable soybean are larger, sweeter and softer than grain soybean. Due to its shorter crop duration, vegetable soybean can fit precisely into the narrow window in crop rotation (Sam et al. 2012).

Improvement for seed weight in soybean by conventional breeding is highly influenced by genotype x environment interaction, it is time consuming and labour intensive and hence, selection for high and stable seed weight requires evaluation in multiple environments over the locations. Molecular markers offer a faster and more accurate approach to breeding for traits such as seed weight, as selection can be based on genotype rather than solely on phenotype. The use of molecular markers for indirect selection of important agronomic traits, or marker-assisted selection (MAS) can improve the efficiency of traditional plant breeding (Allen 1994). This information will also be useful for screening segregating lines in off-season to identify suitable plants for field evaluation. During off-season (in polyhouse), plants normally do not express their true phenotype, therefore molecular data will help in identifying desirable genotypes and will help to accelerate the breeding process. The heritability of seed weight in soybean ranging from 44 to 94 % (Mian et al. 1996) and it is controlled by many genes with additive effects (Brim and Cockerham 1961). The availability of soybean linkage maps (Song et al. 2004; Hyten et al. 2010) has helped immensely in the identification of agronomic trait loci, including quantitative trait loci (QTLs) for seed weight (Teng et

*Corresponding author's e-mail: goruvkhosla@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

al. 2004) during the development of soybean. Several such loci for 100-seed weight have been identified with markers explaining up to 50 % phenotypic variation (Xu et al. 2011; Hu et al. 2013). So far, a few reports on QTL for seed size in vegetable soybean are available. Present study was, therefore, conducted to map QTL for seed weight in a cross between grain and vegetable soybean.

A F_2 mapping population was developed in *kharif* 2014 from a cross between a Punjab state released variety SL 958 with medium seed size (14.3g) and AGS 456, a bold seeded (100-seed weight 27. g) exotic line of vegetable type soybean from Taiwan. The F_1 was grown during *kharif* 2015 and F_2 population comprising of 200 plants was raised in net house during *kharif* 2016. Hundred seed weight of all the F_2 plants and the parents was recorded using electronic precision balance with accuracy up to 0.001g. Data for 100-seed weight was tested for normality (Shapiro-Wilk test) using PAST software.

For genotyping, young leaves were collected from parents and 200 F_2 plants. Genomic DNA of these plants was isolated using the CTAB (Cetyltrimethyl ammonium bromide) method and molecular marker analysis was done following the standard procedure (Allen et al. 2006). A set of 207 SSR markers were selected from linkage map developed by Cregan et al. (1999), covering all the chromosomes. SSR primers were synthesized by IDT Company (India). As per the procedure (Michelmore et al. 1991), two bulks were prepared using ten lines with least seed weight (100-seed weight up to 14.0 g) and ten lines with highest 100 seed weight (above 25.0 g) for the bulked segregant analysis. Primers showing polymorphism between parents were used to amplify bulked DNA and primers polymorphic in the two extreme bulks were identified. Based on the evaluations of DNA bulks, individual F_2 plants were analyzed with co- segregating primers to confirm SSR markers linked to 100-seed weight.

Genotypic and phenotypic data of seed size was analyzed for QTL studies by using QTL Cartographer v. 2.5 software (Wang et al. 2007). Data was first analyzed to recognize markers related with variation for seed weight using single marker analysis (SMA) by all linked and unlinked loci at a statistical threshold of $p < 0.01$. After this, QTL Cartographer's module Zmapqtl (Model 6) was used for composite interval mapping (CIM). QTLs were confirmed by threshold LOD scores based on 1000 permutations for $p < 0.01$. The QTL position, LOD score, the per cent phenotypic variation and additive effect were estimated by CIM for each QTL.

The 100-seed weight of F_2 plants ranged from 8.0 to 27.2g with a mean of 20.1g (Fig. 1). Transgressive segregants were recovered in the F_2 but most of them were for low seed weight. The variation for seed size in F_2 plants followed a normal curve as indicated by significant value (0.07) of Shapiro-Wilk normality test. Near zero values of kurtosis (-0.07) and skewness (-0.17) shows normal distribution of 100-seed weight indicating quantitative nature of genetic control for seed size. The mean for 100-seed weight of the F_2 population (20.1 g) was very close to the mid parent value (20.7 g) with 101 plants having seed weight less than the mean value and 99 plants having seed weight higher than the mean value. Brim and Cockerham (1961) using two crosses, N48-4860 x Lee and Roanoke x Lee of soybean also reported that the mean seed size of population regressed towards mid parent value. Transgressive segregants were found in the direction of lower seed weight only. In a recent study on QTL mapping in a RIL population of soybean conducted by Yashpal et al. (2019), the transgressive segregation was recorded for many traits except 100-seed weight. These observations indicated that a maximum threshold level for improvement of seed weight might have been achieved in soybean.

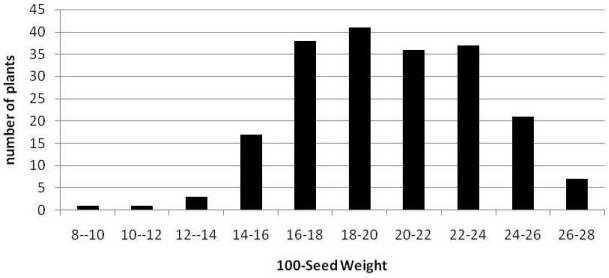

Fig. 1. Distribution for 100-seed weight in F_2 plants

In most of the QTL mapping studies in soybean, the whole population was screened with polymorphic markers, out of which, very few markers showed association with the trait in question as reported in rice also (Gomez et al. 2010). To reduce the efforts and costs associated with genotyping of large mapping populations, for QTL mapping bulked segregant analysis (BSA) was used. In the recent past, BSA has been applied for quantitative traits in various crops such as rice (Venuprasad et al. 2011; Zhang et al. 2009), wheat and maize. Precise application of BSA

depends on the classification of mapping population into distinct classes that roughly reflect the genotypes of target QTLs. This will be more effective if phenotype is a good indicator of the QTL genotype. In present study, a total of 207 SSR primer pairs spanning over 20 chromosomes of soybean genome were used to survey the parental polymorphism between parents. Out of 207, 90 SSR markers were polymorphic, which were further utilized in BSA. Of the 90 polymorphic markers, 18 were polymorphic for the bulks also. Table 1 presents the name and position of these 18 SSR polymorphic markers on nine linkage groups as given by Song et al. (2004).

Table 1. A list of primers polymorphic for bulks for seed size

S.No.	Primer	Linkage group	Position on	$\chi^2_{1:2:1}$	S.No.	Primer	Linkage group	Position on	$\chi^2_{1:2:1}$
1	Sct_034	B2 (14)	51.45	3.51	10	Satt389	D2 (17)	79.23	4.77
2	Satt066	B2 (14)	78.84	1.45	11	Sct_189	I (20)	113.77	4.93
3	Satt560	B2 (14)	97.92	1.92	12	Satt279	H (12)	68.50	1.51
4	Satt357	C2 (6)	151.91	1.64	13	Sat_245	L (19)	115.07	4.34
5	Satt266	D1b (2)	59.61	2.24	14	Sat_244	M (7)	48.86	2.47
6	Satt005	D1b (2)	75.29	3.50	15	Satt175	M (7)	61.93	3.89
7	Satt041	D1b (2)	84.04	2.23	16	Satt336	M (7)	133.83	5.09
8	Sat_069	D1b (2)	102.60	2.59	17	Sat_321	O (10)	19.44	1.51
9	Sat_183	D1b (2)	112.63	1.96	18	Sat_242	O (10)	74.05	2.94

Chr.= Chromosome

Prior to QTL analysis, χ^2 square test was applied to calculate the segregation ratios of all the 18 polymorphic markers. The calculated χ^2 values were then compared with table values at 5 % and 1 % probability levels at 2 degrees of freedom. The segregation behaviour of markers for 202 F_2 progenies was compared with the expected ratio of 1A:2H:1B (1 homozygote from parent P1: 2 hetrozygotes : 1 homozygote from P2). The calculated χ^2 values using observed values of A: H : B and its expected frequency for all the marker loci are presented in the Table 1. All the primer pairs were segregating in normal fashion which was evident from non-significant χ^2 values for all the primer pairs.

Single marker analysis for seed size was estimated using genotypic data of 18 SSR markers and phenotypic data from each of the F_2 individuals. Markers, Satt279 (LG H), Satt266, Satt041, Sat_069, Sat_183 (LG D1b), Satt357 (LG C2), Sct_189 (LG I), Satt389 (LG D2), Sat_244, Satt175 (LG M), Sat_321 and Sat_242 (LG O) explained a significant phenotypic variation with major effect. One major QTL (*qSW-1*) was identified on LG M with Sat_244 and Satt175 as flanking markers having an estimated 19.0 % phenotypic variance with LOD score 9.94 (Table 2).

This QTL was at a distance of 8cM from Sat_244 and 5.04cM from Satt175. Teng et al. (2009) also reported a major effect QTL for seed weight located on 18.58 to 56.29cM on LG M. Two minor QTLs (*qSW-2* and *qSW-3*) were also identified on LG D1b. One QTL (*qSW-2*) was identified in interval Satt041-Sat_069 with an estimated phenotypic variation (R^2) of 6.0 %

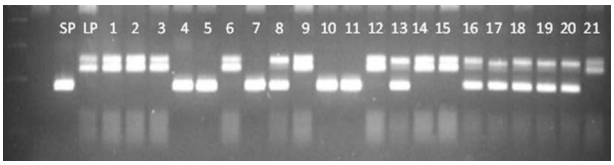

Fig. 2. Segregation of SSR marker Sat_244 in F_2 population (SP=Small seeded parent; LP= Large seeded parent and Lanes 1-21= F_2 plants

with LOD score 4.88 and the other (*qSW-3*) in interval Sat_069-Sat_0183 which was estimating 7.0 % phenotypic variance having a LOD score of 4.95. In a similar study on seed weight in soybean, Kato et al. (2014) reported a QTL for seed weight associated with marker Sat_183 that was explaining 7.3 % phenotypic variation. Bulked segregant analysis, combined with selective genotyping, permitted identification of one major and two minor QTL for seed weight in soybean by reducing the degree of effort needed to identify

Table 2. QTL information regarding seed weight in F_2 population using CIM

S.No.	LG	Marker	Position	LOD	Additive score	R^2
1	D1b	Satt041-Sat_069	31	4.88	1.43	0.06
2	D1b	Sat_069-Sat_0183	41	4.95	1.49	0.07
3	M	Sat_244-Satt175	8	9.94	2.39	0.19

associations between markers and phenotypes. These identified QTLs can be fine mapped and validated for their application in marker assisted breeding for improving seed weight in soybean.

Authors' contribution

Conceptualization of research (GK, BSG); Designing of the experiments (GK, BSG, AS); Contribution of experimental materials (GK, BSG, AS); Execution of field/lab experiments and data collection (GK, PS, AS); Analysis of data and interpretation (GK, BSG, ID); Preparation of the manuscript (GK, BSG, ID).

Declaration

The authors declare no conflict of interest.

References

Allen G. C., Flores-Vergara M. A., Krasynanski S., Kumar S. and Thompson W. F. 2006. A modified protocol for rapid DNA isolation from plant tissues using Cetyltrimethyl-ammonium bromide. Nature Protocols, **1**: 2320-2325.

Allen F. L. 1994. Usefulness of plant genome mapping to plant breeding. *In*: Plant Genome Analysis (ed. Gresshoff P.). CRC Press: Boca Raton, pp 11-18.

Brim C. A. and Cockerham C. C. 1961. Inheritance of quantitative characters in soybeans. Crop Sci., **1**: 187-190.

Cregan P. B., Jarvik T., Bush A. L., Shoemarker R. L., Lark K. G., Kahler A. L., Kaya N., Vantoai T. T., Lohnes D. G., Chung J. and Specht J. E. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci., **39**: 1464-1470.

Cui C. X. and Xuan Y. N. 2007. Analysis of the factors affecting soybean trade in China and suggestions of strategies. World Agric., **29**: 7-10.

Gomez S. M., Boopathi N. M., Kumar S. S., Ramasubramanian T., Chengsong Z., Jeyaprakash P., Senthil A. and Babu R. C. 2010. Molecular mapping and location of QTLs for drought-resistance traits in indica rice (*Oryza sativa* L.) lines adapted to target environments. Acta Physiol. Plant., **32**: 355-364.

Han Y., Li D., Zhu D., Li H., Li X., Teng W. and Li W. 2012. QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. Theor. Appl. Genet., **125**: 671-683.

Hu Z., Zhang H., Kan G., Ma D., Zhang D., Shi G., Hong D., Zhang G. and Yu D. 2013. Determination of the genetic architecture of seed size and shape via linkage and association analysis in soybean (*Glycine max* L. Merr.). Genetica, **141**: 247-254.

Hyten D. L., Choi I-Y., Song Q., Specht J. E., Carter T. E., Shoemaker R. C., Hwang E-Y., Matukumalli L. K. and Cregan P. B. 2010. A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci., **50**: 960-968.

Kato S., Sayama T., Fujii K., Yumoto S., Kono Y., Hwang T. Y., Kikuchi A., Takada Y., Tanaka Y., Shiraiwa T. and Ishimoto M. 2014. A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. Theor. Appl. Genet., **127:** 1365-1374.

Liang H. Z., Li W. D., Wang H. and Fang X. J. 2005. Genetic effects on seed traits in soybean. Acta Genet. Sin., **32**: 1199-1204.

Mian M. A. R., Bailey M. A., Tamulonis J. P., Shipe E. R., Carter J. T. E., Parrott W. A., Ashley 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.

Michelmore R. W., Paranand I. and Kessele R. V. 1991. Identification of markers linked to disease resistance genes by bulk segregant analysis: A rapid method to detect markers in specific genome using segregant population. Proc. Nat. Acad. Sci. USA, **88**: 9828-9832.

Sam C. E., Pantalone V. R., Kopsell D. A., Zivanovic S. and Deyton D. E. 2012. Edamame: a potential high value crop for growers. In: Proceedings of the Mid-Atlantic Fruit and Vegetable Convention. Mid-Atlantic Fruit and Vegetable Convention, Hershey, PA, USA.

Song Q. J., Marek L. F., Shoemaker R. C., Lark K. G., Concibido V. C., Delannay X., Specht J. E. and Cregan P. B. 2004. A new integrated genetic linkage map of the soybean. Theor. Appl. Genet., **109**: 122-128.

Teng W., Han Y., Du Y., Sun D., Zhang Z., Qiu L., Sun G. and Li W. 2009. QTL analyses of seed weight during the development of soybean (*Glycine max* L. Merr.). Heredity, **102:** 372-380.

Venuprasad R., Bool M. E., Quiatchon L., Sta Cruz M. T., Amante M. and Atlin G. N. 2011. A large-effect QTL for rice grain yield under upland drought stress on chromosome 1. Mol. Breed. http://dx.doi.org/10. 1007/

s11032-011-9642-2.

Wang S., Basten C. J. and Zeng Z. B. 2007. Windows QTL cartographer 2.5. Department of statistics, North Carolina State University, Raleigh, NC, USA. [2012-04-23]. http://statgen.ncsu.edu/qtlcart/WQTLCart.html.

Xu Y., Li H. N., Li G. J., Wang X., Cheng L. G. and Zhang Y. M. 2011. Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* L. Merr.). Theor. Appl. Genet., **122**: 581-594.

Yashpal, Rathod D. R., Chandra S., Kumar Anil, Yadav R. R. and Talukdar A. 2019. Deploying inter-specific recombinant inbred lines to map QTLs for yield-related traits in soybean Indian J. Genet., **79**(4): 693-703. DOI: 10.31742/IJGPB.79.4.7.

Zhang B., Chen P., Florez-Palacios S. L., Shi A., Hou A. and Ishibashi T. 2010. Seed quality attributes of food-grade soybeans from the US and Asia. Euphytica, **173**: 387-396.

Zhang G. L., Chen L. Y., Xiao G. Y., Xiao Y. H., Chen X. B. and Zhang S. T. 2009. Bulked segregant analysis to detect QTL related to heat tolerance in rice (*Oryza sativa* L.) using SSR markers. Agric. Sci. China, **8**: 482-487.