High-density linkage map reveals QTL for Type-I seed coat cracking in RIL population of soybean [*Glycine max* (L.) Merr.]

Beom Kyu Kang · **Jeong Hyun Seo** · **Hyun Jo** · **Krishnanand P. Kulkarni** ·
Man Soo Choi · **Hyun Tae Kim** · **Jeong Dong Lee** · **Sanjeev K. Dhungana** ·
Hong Sik Kim · **Jae Hyeon Oh** · **Ji Hee Park** · **Sang Ouk Shin** · **In Youl Baek**

Abstract Seed coat cracking (SCC), particularly the Type-I irregular cracking, is critical in determining the quality of appearance and commercial value of soybean seeds. The objective of this study was to identify the quantitative trait loci (QTLs) for SCC with high-density genetic map. One hundred sixty-seven recombinant inbred lines (RILs) developed from a cross between Uram (SCC-resistant) and Chamol (SCC-susceptible) were evaluated for SCC over 2 years (2016–2017). The QTL analysis identified 12 QTLs located on chromosomes 2 (D1b), 6 (C2), 8 (A2), 9 (K), 10 (O), 12 (H), 19 (L), and 20 (I). Out of the 12 QTLs, *qSC2-1*, *qSC9, SC10-1*, *qSC10-2*, and *qSC12* were novel QTLs and the other seven QTLs (*qSC2-2*, *qSC2-3*, *qSC6*, *qSC8*, *qSC19-1*, *qSC19-2*, and *qSC20*) were found to co-localize with the previously identified QTLs. The mean SCC of the RILs of early maturity group was significantly higher than that of the late maturity group, suggesting an association between SCC and maturity loci. In addition, although 10 QTLs were distantly located from the maturity loci (*E1*, *E3*, *E4*, *E7*, and *E10*), *qSC10-1* and *qSC10-2* co-localized with the maturity loci *E2*. The results obtained in this study provide useful

B. K. Kang (✉) · J. H. Seo · H. T. Kim ·
S. K. Dhungana · H. S. Kim · J. H. Park ·
S. O. Shin · I. Y. Baek
Upland Crop Breeding Research Division, Department of Southern Area Crop Science, National Institute of Crop Science (NICS), Rural Development Administration (RDA), Miryang 50424, Korea
e-mail: hellobk01@korea.kr

H. Jo · J. D. Lee
School of Applied Biosciences, Kyungpook National University (KNU), Daegu 41566, Korea

K. P. Kulkarni
Department of Agriculture and Natural Resources, College of Agriculture, Science and Technology, Delaware State University, Dover, DE 19901, USA

M. S. Choi
Crop Foundation Research Division, National Institute of Crop Science (NICS), Rural Development Administration (RDA), Jeonju 55365, Korea

J. H. Oh
Gene Engineering Division, Department of Agricultural Biotechnology, National Institute of Agricultural Sciences (NAAS), Rural Development Administration (RDA), Jeonju 54874, Korea

genetic information on SCC which could be used in the SCC breeding programs.

Introduction

Soybean [*Glycine max* (L.) Merr.] is one of the major field crops cultivated globally. Because of plentiful protein and oil contents in soybean seed, it is used for diverse purposes such as food, feed, fuel, and other industrial usages (Masuda and Goldsmith 2009). In a few Asian countries including Korea, several whole-seed-based soybean food recipes are popular and have been a part of traditional foods. Therefore, the quality of soybean seed appearance is considered as an important factor for commercial value.

Seed coat cracking (SCC) is one of the critical traits in determining the visual quality of seed. The SCC can induce and increase the possibility of splitting, damaging, and pathogen infection of the seed. Also, SCC decreases seed germination and emergence when seeds are planted (Yaklich and Barla-Szabo 1993). The SCC can be classified into two types: Type-I is the irregular cracking on seed coat, whereas Type-II is the net-like cracking on seed coat (Liu 1949). Type-II seeds are produced and sold, sometimes, in local markets because of the unique seed coat patterns, while Type-I seeds have a significantly decreased commercial value due to the irregular cracking.

The Type-I cracking results from the separation of the epidermal (palisade cells) and hypodermal (hourglass cells) tissues, which exposes the underlying parenchyma tissue (Yaklich and Barla-Szabo 1993). The SCC may be induced by exposure to chilling temperature (10–18 °C) at the flowering stage (Takahashi 1997). In previous studies, *I* (responsible for the distribution of seed coat color), *T* (responsible for pubescence and seed coat color), and *E1* and *E5* (responsible for flowering and maturity) loci are found to suppress the SCC at low temperatures (Takahashi 1997; Takahashi and Abe 1999), whereas *E2* and *T* loci are found to induce the SCC in pods-removing treatment (Yang et al. 2002).

To evaluate the SCC of different soybean genotypes, the SCC is promoted by using artificial methods, such as pod-removal, drying of imbibed seeds, and application of an ethychlozate (ethylene generating reagent) (Yang et al. 2002). The conventional approaches for screening SCC resistant lines are time-consuming and labor-intensive due to multiple steps involved in the evaluation and complicated genetic backgrounds as well as the existence of an interaction between genetic and environmental effects (Ha et al. 2012). Recent advances in the sequencing and genotyping technologies have facilitated genetic study for many complex traits such as seed fat, protein, seed size, and seed starch content in soybean (Ha et al. 2014; Asekova et al. 2016; Dhungana et al. 2017; Kulkarni et al. 2016, 2018). For SCC, Oyoo et al. (2010) identified two QTLs, *cr1* on chromosome 2 (D1b) and *cr2* on chromosome 7 (M), using a mapping population of 95 recombinant inbred lines (RILs) genotyped with 1015 simple sequence repeat (SSR) markers. In another study, Ha et al. (2012) studied QTL, epistatic effects, and QTL-by-environment interactions for SCC in a 117 RILs population genotyped with 138 SSR markers, and identified 10 QTLs. Out of the 10 QTLs, three QTLs (*qSCC2-1*, *qSCC9*, and *qSCC20*) were identified in more than two environments. Saruta et al. (2019) identified the QTL *qScr20-1* on chromosome 20 (I) using 172 RILs genotyped with 264 SSR markers.

For a comprehensive understanding of the genetic basis of SCC in soybean, it is necessary to identify QTLs using different genetic background across various environments. In the present study, we evaluated a mapping population comprising of 167 RILs across two environments, and identified QTLs associated with SCC using a high-density linkage map constructed by 5179 SNP markers (Kang et al., unpublished). The investigation of QTLs and phenotypic variation can expand knowledge for SCC, Type-I irregular cracking, in soybean.

Materials and methods

Plant materials and growing conditions

A mapping population comprised of 167 RILs, derived from a cross between SCC-resistant Uram (Ko et al. 2016) and SCC-susceptible Chamol (Ko et al. 2018), was developed from 2012 to 2017. Figure 1 shows the appearance of irregular cracking and normal seed coat
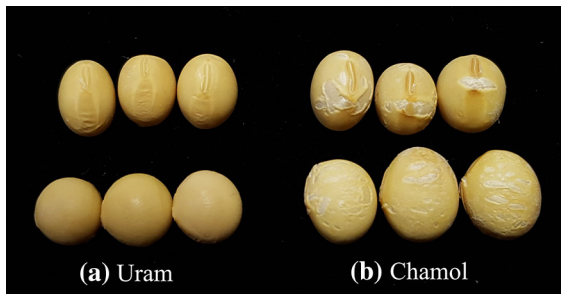
Fig. 1 The seed appearance of normal and Type-I irregular seed coat cracking in soybean from parental cultivars. **a** Normal seed of female parent, Uram, and **b** Type-I irregular cracked seed of male parent, Chamol

of Chamol and Uram. Uram is a late-maturing, whereas Chamol is an early-maturing cultivar. Uram grows taller with a higher-positioned first pod than Chamol. However, both parental cultivars have white pubescence. In 2012, the female parent Uram was crossed with the male parent Chamol. The F_1 seeds were planted in Daegu Experiment Station, NICS, RDA (35° 90′ N 128° 44′ E, Korea) in 2013. In the subsequent year (2014), F_2 population was planted in the same location. One hundred sixty-seven plants derived from the F_2 population were advanced from F_3 to F_5 through single seed descent method in Hung Loc Agricultural Center (10° 56′ N 107° 04′ E, Vietnam) in 2015. The $F_{5:6}$ RILs were planted in Daegu Experiment Station over 2 years (2016 and 2017) in a randomized block design with two blocks. Planting dates were June 28th in 2016 and June 29th in 2017. The RILs were grown in the black vinyl-mulched 2 m long rows those spaced 60 cm apart. Seeds were sown manually keeping 15 cm between hills, and plants were thinned to keep one seedling per hill. Compost (10 ton ha^{-1}) and chemical fertilizers (N–P–K: 30–30–34 kg ha^{-1}) were applied during field preparation.

Evaluation of seed coat cracking

The 167 RILs and parents planted in 2016 and 2017 were harvested at maturity and evaluated for SCC. One hundred seeds were randomly collected in triplicate from each plot, the number of irregularly cracked (Type-I) seed was counted, and expressed as percentage seed cracked.

Statistical analysis

Analysis of variance (ANOVA) was conducted and frequency distribution was obtained using R Studio (Ver 1.1.419). The descriptive statistical parameters (mean, minimum, maximum, median, standard deviation (SD), variance (VAR), coefficient of variation (CV), kurtosis, and skewness) were generated using Microsoft Excel 2016. The environment, genotype, and their interaction were considered as a fixed effect, and the broad-sense heritability (h^2) was estimated from ANOVA using the following formula:

$$h^2 = \sigma_g^2/\sigma_p^2, h^2 = \sigma_g^2/\sigma_g^2 + \left(\sigma_{gy}^2/y\right) + \left(\sigma_e^2/ry\right),$$

where 'y', 'g' and 'r' are number of year, genotype, and replication, respectively; σ_g^2, σ_{gy}^2, and σ_e^2 are components of variance for genotypes, interaction between genotype and environment, and error, respectively (Toker 2004; Kulkarni et al. 2017).

Linkage mapping and QTL analysis

Young trifoliate leaves from single plant of F_6 line derived from F_5 plant were collected and used for DNA extraction. The DNA was extracted using QIAGEN DNeasy® plant mini kit (Qiagen Sciences Inc., Germantown, MD, USA). The extracted DNA was genotyped with 180 K AXIOM® SoyaSNP array (Lee et al. 2015) and scanned with a GeneTitan® Scanner (Affymetrix, Santa Clara, CA, USA).

The genetic linkage map was constructed with 180,375 genome-assigned SNPs, excluding 586 scaffolds region in the whole 180,961 SNPs. A total of 20,046 SNPs showed polymorphism between parental cultivars. The low polymorphism found between the parental lines might be due to the reduced genetic diversity existed among soybean cultivars that resulted as a consequence of domestication and development of commercial varieties (Li et al. 2013; Achard et al. 2020). The genetic map construction and QTL analysis were performed using the polymorphic markers in QTL IciMapping Ver. 4.1 (Meng et al. 2015; Wang et al. 2016). The polymorphic markers were subjected to the Binning function of IciMapping considering a missing rate (5%) and segregation distortion ($P < 0.001$). The mapping options were set as follows: 3.0 LOD (logarithm of odds) grouping, 'nnTwoOpt' ordering, and five size of window for sum of adjacent recombinant frequencies (SARF).

Kosambi's mapping function was used in transforming recombination frequencies into centimorgan (cM) distances (Kosambi 1943).

The QTLs were detected using inclusive composite interval mapping of additive QTLs (ICIM-ADD) with parameters of 1.0 step and 1,000 permutation tests at $P \leq 0.05$ (Li et al. 2007). The figure of linkage maps showing QTL positions was constructed using Map-Chart 2.32 (Voorrips 2002).

The QTL for SCC identified in this study was named by combining different letters and numbers: q; quantitative trait locus, SC; seed coat cracking; the numbers followed by the letters indicate the chromosome harboring the QTL. Thus, qSC2-1 and qSC6, respectively, denote the first QTL on chromosome 2 and the single QTL on chromosome 6.

Results

ANOVA and phenotypic analysis

The SCC of the parental cultivars and the RIL population was measured in two-year environments, and ANOVA was used to analyze genotype, environment, and genotype by environment interaction (G × E) effects on the SCC variation (Table 1). Genotype, environment and G × E effects were significant for the SCC ($P < 0.001$). The estimated broad-sense heritability of SCC was 81.5% which suggested that the higher proportion of variation for SCC was due to the genetic effects was more than the environmental effects.

The descriptive statistics of the SCC variation in the RIL population are given in Table 2. The SCC of Uram and Chamol were 2.0% and 36.5% in 2016, 0.2% and 21.2% in 2017, and 1.1% and 28.8% in combined year (mean of 2016 and 2017), respectively. The mean, minimum, maximum, and median value of RILs were 18.8%, 0.3%, 77.5%, and 13.2% in 2016, 7.3%, 0.0%, 54.8%, and 2.2% in 2017, and 12.9%, 0.3%, 65.3%, and 7.4% in combined year. The SD, VAR, and CV of RILs were 17.3, 299.5, and 92.0% in 2016, 11.3, 127.7, and 154.2% in 2017, and 13.3, 176.4, and 102.6% in combined year.

The skewness values of RILs in 2016, 2017, and combined year were more than 0 (Table 2), and the phenotypic distribution of the SCC in RILs was right-skewed (right-tailed) in all the environments (Fig. 2). The kurtosis values of RILs in 2016 and the combined year were less than 3, but the value in 2017 was more than 3 (Table 2). It indicated that the phenotypic distribution of RILs was less peaked than normal distribution in 2016 and combined year, but more peaked in 2017 (Fig. 2).

The SCC variation in the combined year was different by the group of maturity days (MD, from seeding to maturity) (Table 3). In early maturity group (MD ≤ 100 days, n = 48), the SCC values were 20.0% of mean value, 0.4% of minimum value, 65.3% of maximum value, and 15.4 of SD. The SCC values for normal maturity group (100 < MD ≤ 110 days, n = 86) were 12.1% of mean value, 0.3% of minimum value, 49.8% of maximum value, and 12.1 of SD. The SCC values for late-maturity group (MD > 110 days, n = 33) were 4.3% of mean value,

Table 1 Analysis of variance for environments (E), genotypes (G) and G × E interaction for Type-I seed coat cracking of soybean RIL population evaluated in 2016 and 2017

SV[a]	DF[b]	SS[c]	MS[d]	F-value	P value	h^2 (%)[e]
Ennvironment (E;Year)	1	62,066	62,066	1058.1	< 0.001	81.5
Genotype (G)	166	346,312	2086	35.6	< 0.001	
G × E	166	63,999	386	6.6	< 0.001	
Residuals	1630	95,611	59			

[a]Single value

[b]Degree of freedom

[c]Sum of square

[d]Mean square

[e]Broad-sense heritability

Table 2 The result of phenotypic evaluation for Type-I seed coat cracking for parental cultivars and 167 RILs cultivated in 2016, 2017, and combined years

Year	Parental cultivars		RIL[a] (n = 167)								
	Uram (P1) (%)	Chamol (P2) (%)	Mean (%)	Min (%)[b]	Max (%)[c]	Med[d]	SD[e]	Var[f]	CV (%)[g]	Kurtosis	Skewness
2016	2.0	36.5	18.8	0.3	77.5	13.2	17.3	299.5	92.0	0.8	1.2
2017	0.2	21.2	7.3	0.0	54.8	2.2	11.3	127.7	154.2	5.0	2.3
Combined	1.1	28.8	12.9	0.3	65.3	7.4	13.3	176.4	102.6	2.5	1.6

[a]Recombinant inbred lines

[b]Minimum value

[c]Maximum value

[d]Median value

[e]Standard variation

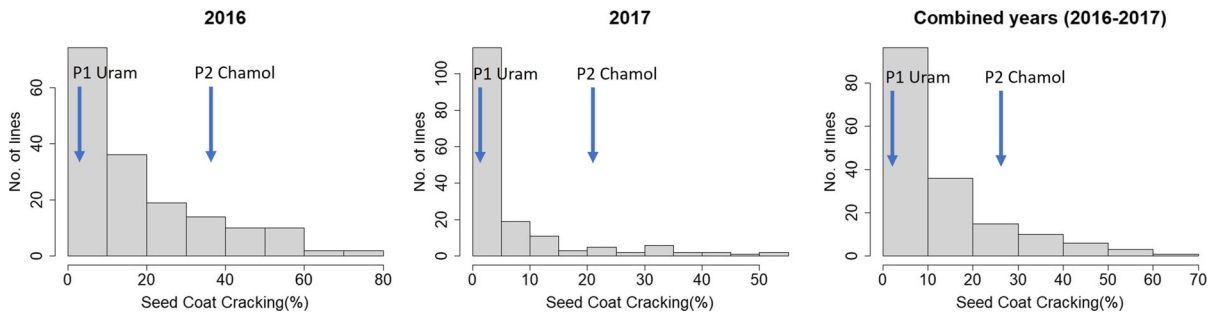[f]Variance

[g]Coefficient of variation

Fig. 2 The phenotypic distribution of Type-I seed coat cracking of RILs and parental cultivars evaluated in 2016, 2017, and combined year

Table 3 The mean, standard deviation, minimum, and maximum value for Type-I seed coat cracking of 167 RILs grouped by maturity days in combined year (2016 – 2017)

Group (n = 167)	Seed coat cracking			
	Mean (%)	Min[b]	Max[c]	SD[d]
Early (MD[a] ≤ 100, n = 48)	20.0a[e]	0.4	65.3	15.4
Normal (100 < MD ≤ 110, n = 86)	12.1b	0.3	49.8	12.1
Late (MD > 110, n = 33)	4.8c	0.3	18.0	4.3

Same characters in the column were not significantly different by DMRT at 0.05 level of probability

[a]Maturity days: days of seeding to maturing

[b]Minimum value

[c]Maximum value

[d]Standard variation

0.3% of minimum value, 18.0% of maximum value, and 4.3 of SD. The correlation coefficient between SCC and MD was − 0.43 ($P < 0.001$) (data not shown).

Linkage mapping and QTL analysis

The 167 RILs and both parents were genotyped by using 180,375 SNPs, out of which 20,046 SNPs were polymorphic between the parents. After binning (missing rate 5% and segregation distortion $P < 0.001$), a total of 5179 SNP markers remained and were used for linkage map construction. The average number of markers across 20 linkage groups was 259 SNPs, and the average distance between the SNPs was 0.7 cM. The total map length of 20 linkage groups spanned 2758 cM and averaged 138 cM for each linkage group. The smallest linkage group formed for chromosome 16 (J), whereas the largest linkage group formed for chromosome 2 (D1b) (Table S1).

The 12 QTLs associated with SCC were identified on 8 chromosomes (Table 4 and Fig. 3). The ICIM analysis detected significant QTL regions for SCC in two-year environments and combined year data. In 2016, four QTLs (*qSC2-1, qSC10-1, qSC12* and *qSC19-1*) were detected on chromosomes 2 (D1b), 10 (O), 12 (H) and 19 (L), respectively. These QTLs explained 7.3% – 15.6% of phenotypic variance explained (PVE) with the LOD score ranging from 4.9 to 8.9. In 2017, five QTLs (*qSC2-2, qSC6, qSC8, qSC9, and qSC10-2*) were detected on chromosomes 2 (D1b), 6 (C2), 8 (A2), 9 (K) and 10 (O) explaining 4.3% – 16.5% of PVE at LOD score range of 4.3 – 16.6. In combined year, six QTLs (*qSC2-3, qSC6, qSC10-2, qSC12, qSC19-2,* and *qSC20*) explaining 4.1–12.9% of PVE and 4.1–12.9 of LOD scores were detected on chromosome 2 (D1b), 6 (C2), 10 (O), 12 (H), 19 (K) and 20 (I). The highest and lowest LOD scores were found in *qSC8* and *qSC19-2*, respectively. Similarly, QTL *qSC8* and *qSC9* contributed for the highest and lowest PVE, respectively.

Discussion

The SCC of soybean, especially Type-I irregular cracking, is an important phenotype in determining the commercial value of seeds. The main purpose of this study was to identify QTL for SCC using a biparental mapping population. The results of phenotypic evaluation indicated that the SCC was significantly influenced by genotype, environment, and their interaction. The high value of the broad-sense heritability showed that the genotypic factor was more influential than environmental factor in determining the SCC variation. When the heritability is higher than 50%, the target quantitative phenotype can be considered as a selection marker for subsequent generations, considering the trait variation is mainly based on genetic inheritance. The SCC of RIL population showed transgressive segregation, especially over the susceptible parent because the resistant parent showed small variance (2.33), whereas the susceptible parent showed large variance (49.92). Similar results of right-skewed distribution were also found in previous studies (Oyoo et al. 2010; Ha et al. 2012; Saruta et al. 2019).

The average distance between SNP markers, in this study, was 0.7 cM, which was relatively of higher density compared to previous QTL studies for SCC (Oyoo et al. 2010; Ha et al. 2012; Saruta et al. 2019). Construction of the high-density linkage map is important for precise mapping of QTLs and their potential application in breeding programs.

The previous studies on SCC suggested that the maturity loci (*E1*, *E2*, and *E5*) and pigment loci (*T* and *I*) were associated with SCC variation in specific environments (Takahashi 1997; Takahashi and Abe 1999; Yang et al. 2002). In the present as well as previous QTL studies, most of the QTLs for SCC were located in the same linkage groups where the maturity loci exist. Therefore, we investigated the SCC variation in the RILs considering their maturity period (early, normal, and late), and found significant differences among the groups (Table 3). The maturity can be an important factor to affect SCC variation. We also compared the physical locations of the QTLs identified in the present study with that of the maturity loci and previously detected QTLs based on the information obtained from SoyBase (https://www.soybase.org/, accessed February 2020).QTLs *qSC2-2* and *qSC2-3* co-localized with *qSCC2-1* (Ha et al. 2012) and *cr1* (Oyoo et al. 2010). The physical location of *qSC6* overlapped that of *qSCC6* which located at about 30 cM from three clustered loci *E1*, *E7*, and *T* (Molnar et al. 2003; Ha et al. 2012). *E1* and *T* were known to suppress the SCC at low temperatures and possibly

Table 4 The QTLs identified for Type-I seed coat cracking evaluated in 2016, 2017, and combined years with the RIL population developed from the cross between cultivars Uram and Chamol

QTL	Year	Chr (LG)[a]	Position (cM)	Left marker	Physical position (bp)	Right marker	Physical position (bp)	LOD[b]	PVE (%)[c]	Add[d]	CI[e]
qSC2-1	2016	2 (D1b)	1	AX-90397999	837,860	AX-90477326	859,912	7.0	10.2	5.7	0–1.5
qSC2-2	2017	2 (D1b)	13	AX-90352789	4,058,781	AX-90328934	4,129,981	11.5	11.4	4.7	12.5–13.5
qSC2-3	Combined[f]	2 (D1b)	15	AX-90405774	4,222,566	AX-90342458	4,369,723	12.9	15.0	5.3	14.5–15.5
qSC6	2017	6 (C2)	60	AX-90457808	1,470,272	AX-90352417	6,368,572	6.1	8.3	4.1	51.5–69.5
qSC6	Combined	6 (C2)	59	AX-90457808	1,470,272	AX-90352417	6,368,572	6.8	11.2	4.6	51.5–67.5
qSC8	2017	8 (A2)	46	AX-90374998	17,185,579	AX-90329948	17,596,226	16.6	17.3	5.8	45.5–46.5
qSC9	2017	9 (K)	111	AX-90430218	42,358,891	AX-90445889	46,182,296	4.3	3.9	2.8	103.5–111.5
qSC10-1	2016	10 (O)	105	AX-90506930	44,501,796	AX-90472723	44,728,112	7.2	10.7	− 5.8	101.5–105.5
qSC10-2	2017	10 (O)	111	AX-90490202	45,818,758	AX-90513989	45,969,241	12.0	11.4	− 4.7	110.5–111.5
qSC10-2	Combined	10 (O)	111	AX-90490202	45,818,758	AX-90513989	45,969,241	10.7	12.2	− 4.8	110.5–111.5
qSC12	2016	12 (H)	39	AX-90336976	7,992,664	AX-90436508	8,085,400	4.9	7.3	4.9	38.5–39.5
qSC12	Combined	12 (H)	39	AX-90336976	7,992,664	AX-90436508	8,085,400	5.1	5.7	3.3	38.5–39.5
qSC19-1	2016	19 (L)	112	AX-90489570	42,347,892	AX-90395600	44,969,419	8.9	15.6	− 7.0	106.5–113.5
qSC19-2	Combined	19 (L)	116	AX-90395600	44,969,419	AX-90472983	45,576,271	4.1	4.8	− 3.0	109.5–120.5
qSC20	Combined	20 (I)	21	AX-90462479	2,949,010	AX-90365081	31,450,550	6.3	6.8	− 3.6	18.5–21.5

[a]Chromosome (linkage group)

[b]Logarithm of odd

[c]Phenotypic variance (%) explained by the QTL

[d]Additive effect; negative additive effects indicate contribution of 'Chamol' for SCC increase

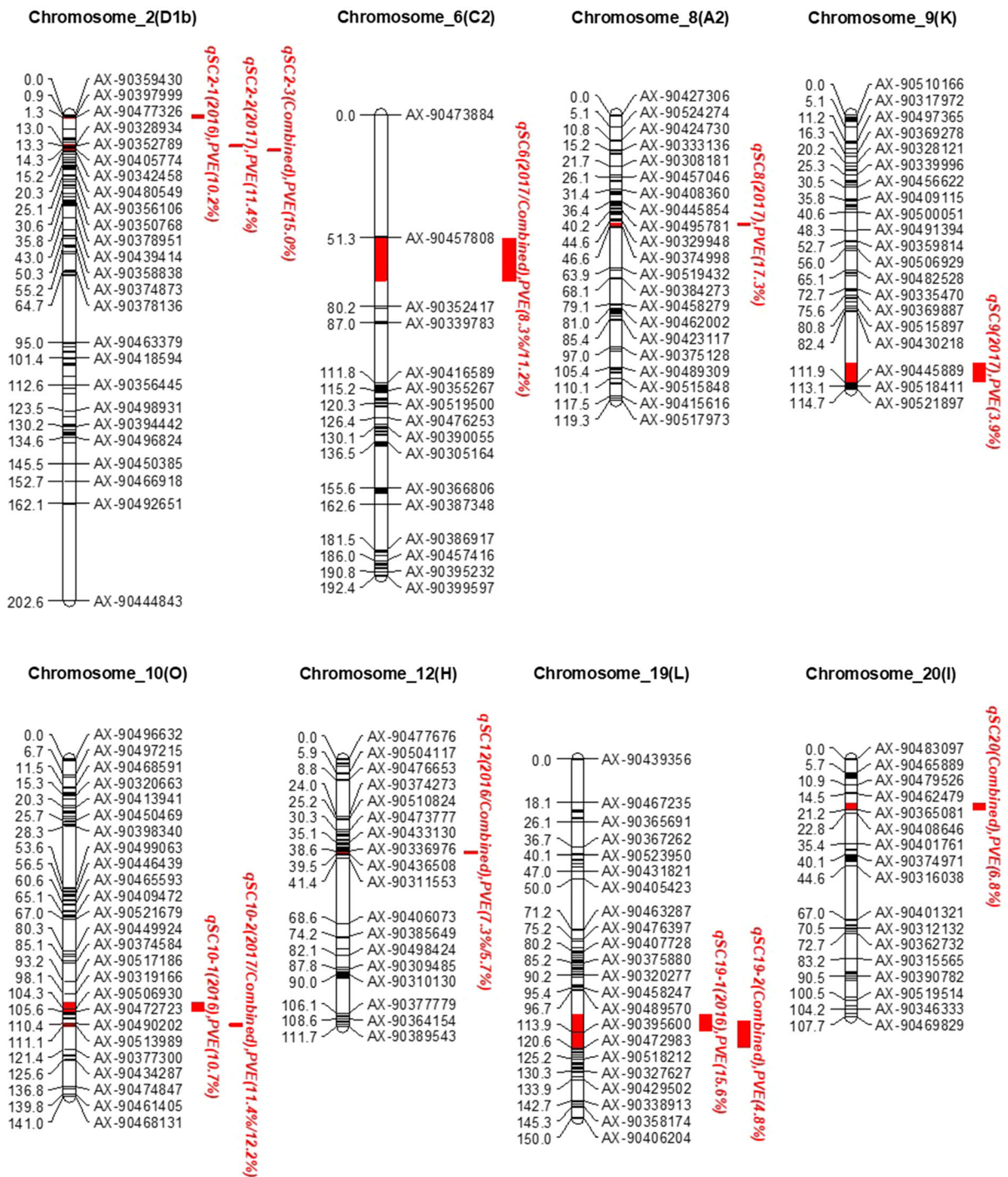[e]Confidence interval

[f]Mean of 2016 and 2017

Fig. 3 Chromosomal locations of the detected QTLs controlling Type-I seed coat crackiing from 167-RIL population crossed between Uram and Chamol, and evaluated 2016, 2017, and combined year, and genotyped with 5179 SNPs. QTLs are marked with bars. The bar length represents the marker interval of each QTL

had roles for controlling SCC variation (Takahashi 1997; Takahashi and Abe 1999). However, the physical position of the markers for *qSC6* was 14.3 Mb, 2.51 Mb, and 1.24 Mb away from the loci *E1*, *E7*, and *T*, respectively (Toda et al. 2002; Molnar et al. 2003; Dissanayaka et al. 2016). Also, the pubescence color, relating to *T* locus, was not different between the parental cultivars as well as among the RILs. All the parents and RILs had white pubescence. Therefore, SCC variation found in this study might not be related to loci *E1, E7*, and *T*. *qSC8* co-localized with *qSCC8* (Ha et al. 2012), and located at 2.8 Mb distance from *E10* locus (Samanfar et al. 2017). *qSC19-1* and *qSC19-2* were found to cover the physical location of *qSCC19* (Ha et al. 2012), and located 2 Mb far from *E3* locus (Mao et al. 2017). Similarly, *qSC20* co-localized with *qSCC20* (Ha et al. 2012; Saruta et al. 2019). These results showed that the SCC variation found in the RIL population was not directly related to the maturity loci *E3* and *E10*, and the association of SCC with *E4* was not also clear (Molnar et al. 2003).

qSC2-1, *qSC9*, *qSC10-1*, *qSC10-2*, and *qSC12* were the novel QTLs for SCC detected in this study. Out of the four chromosomes that harbored the novel QTLs, only chromosome 10 was found to contain the maturity loci. *E2* locus found on chromosome 10 (O) has been reported to induce SCC in one of the treatment groups of pod-removing experiments in soybean (Yang et al. 2002). The chromosomal region between *qSC10-1* and *qSC10-2* covered a *GIGANTEA* ortholog, *GmGIa* gene (*Glyma.10g221500*) that was identified as *E2* locus in soybean genome (Watanabe et al. 2011). On the other three chromosomes (2, 9, and 10), several candidate genes associated with flowering and seed maturing were found in the interval of the novel QTLs *qSC2-1*, *qSC9*, and *qSC12* (Supplementary Table S2). The marker interval of *qSC2-1* includes four genes, out of which *Glyma.02g008300* and *Glyma.02g008400* are related to pectinesterase which affects the accumulation of methanol in maturing soybean seeds (Markovic and Obendorf 2008) and *Glyma.02g008500* is related to protein kinase domain playing important roles in seed maturation of rice and sandalwood (Kawasaki et al. 1993; Anil et al. 2000). A protein kinase domain is also associated with the activity of oil bodies of several plant species, including soybean seed (Anil et al. 2003). The physical location of *qSC9* overlapped the marker *Gm09_43508261* associated with flowering time in

soybean (Mao et al. 2017). The marker interval of *qSC12* includes seven genes, out of them *Glyma.12g095700* is related to seed maturation protein PM37 from NCBI database (https://ncbi.nlm.nih.gov, accessed February 2020).

A few studies suggested that SCC was related to several maturity loci (Takahashi 1997; Takahashi and Abe 1999; Yang et al. 2002), which was also noticed in the RIL population with a higher SCC in the early-mature soybean lines. During the soybean seed coat development, various cells and tissues undergo several changes after fertilization until maturation (Shibles et al 2004). The flowering time and subsequent seed development may vary with genotypes, and are also influenced by the growing environmental conditions such as temperature. Temperature affects cell division (Francis and Barlow 1988) and can induce variation in the physical appearance of soybean seed coat. The scatter plot indicates that the late-mature group has lower SCC than the early-mature group, even though few early-mature lines have low level of SCC (Fig. 4). Thus, SCC varied in the RILs of different groups of maturity, suggesting the effect of maturity loci, especially *E2* locus for the variation in SCC along with maturity in this study. Further research using the population derived from the cross between SCC-resistant and -susceptible lines but without a difference in flowering and maturing time could be useful to investigate the relationship between maturity and SCC. To precisely determine the genetic regions
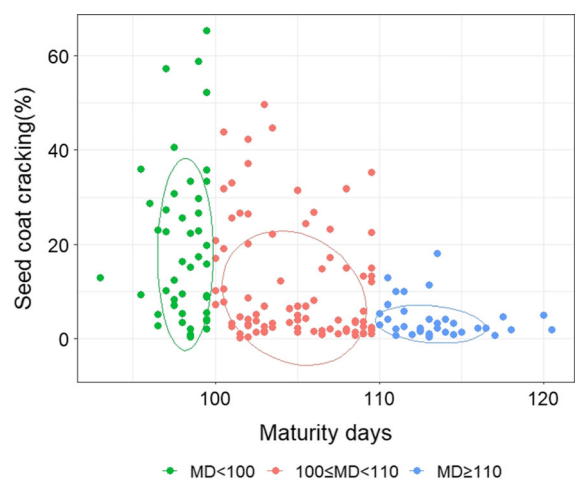

Fig. 4 The scatter plot between seed coat cracking and maturity days of RILs in combined years (2016–2017) and clustered by maturity group. MD indicates maturity days

affecting SCC and develop useful markers, the whole-genome resequencing data of both parents would be required to identify the sequence variations within candidate genes (Asekova et al. 2016). The QTLs for SCC and the potential relation between SCC and maturity identified in this study could provide useful information on the genetic control of SCC in soybean. This information can be of great significance for soybean breeding and development of SCC-resistant cultivar by adopting marker-assisted selection technology.

References

Achard F, Butruille M, Madjarac S, Nelson PT, Duesing J, Laffont JL et al (2020) Single nucleotide polymorphisms facilitate dus testing of soybean cultivars for plant variety protection. Crop Sci. https://doi.org/10.1002/csc2.20201

Anil VS, Harmon AC, Rao KS (2000) Spatio-temporal accumulation and activity of calcium-dependent protein kinases during embryogenesis, seed development, and germination in sandalwood. Plant Physiol 122:1035–1043. https://doi.org/10.1104/pp.122.4.1035

Anil VS, Harmon AC, Rao KS (2003) Temporal association of Ca^{2+}-dependent protein kinase with oil bodies during seed development in *Santalum album* L.: its biochemical characterization and significance. Plant Cell Physiol 44(4):367–376. https://doi.org/10.1093/pcp/pcg046

Asekova S, Kulkarni KP, Patil G, Kim M, Song JT, Nguyen HT, Shannon JG, Lee JD (2016) Genetic analysis of shoot fresh weight in a cross of wild (*G. soja*) and cultivated (*G. max*) soybean. Mol Breed 36:103. https://doi.org/10.1007/s11032-016-0530-7

Dissanayaka A, Rodriguez TO, Yan SDF, Githiri SM, Rodas FR, Abe J, Takahashi R (2016) Quantitative trait locus mapping of soybean maturity gene E5. Breed Sci Prev. https://doi.org/10.1270/jsbbs.15160

Dhungana SK, Kulkarni KP, Park CW, Jo H, Song JT, Shin DH, Lee JD (2017) Mapping quantitative trait loci controlling soybean seed starch content in an interspecific cross of 'Williams 82' (*Glycine max*) and 'PI 366121' (*Glycine soja*). Plant Breed 136(3):379–385. https://doi.org/10.1111/pbr.12480

Francis D, Barlow PW (1988) Temperature and the cell cycle. Symp Soc Exp Biol 42:181–201

Ha BK, Kim HK, Kang ST (2012) Mapping QTLs with epistatic effects and QTL-by-environment interactions for seed coat cracking in soybeans. Euphytica 186:933–942. https://doi.org/10.1007/s10681-012-0719-8

Ha BK, Kim HJ, Velusamy V, Vuong TD, Nguyen HT, Shannon JG, Lee JD (2014) Identification of quantitative trait loci controlling linolenic acid concentration in PI483463 (*Glycine soja*). Theor Appl Genet 127(7):1501–1512. https://doi.org/10.1007/s00122-014-2314-y

Kawasaki T, Hayashida N, Baba T, Shinozaki K, Shimada H (1993) The gene encoding a calcium-dependent protein kinase located near the sbe1 gene encoding starch branching enzyme I is specifically expressed in developing rice seeds. Gene 129(2):183–189. https://doi.org/10.1016/0378-1119(93)90267-7

Ko JM, Han WY, Kim HT, Lee YH, Choi MS, Lee BW et al (2016) Soybean cultivar for soy-paste, 'Uram' with mechanization harvesting, large seed, disease resistance and high yield. Korean J Breed Sci 48:301–306. https://doi.org/10.9787/KJBS.2016.48.3.301

Ko JM, Kim HT, Han WY, Baek IY, Yun HT, Lee YH et al (2018) 'Chamol', and early maturing, high yield, and large-seed soybean cultivar for double cropping. Korean J Breed Sci 50:478–484. https://doi.org/10.9787/KJBS.2018.50.4.478

Kosambi DD (1943) The estimation of map distances from recombination values. Ann Eugenic 12:172–175. https://doi.org/10.1007/978-81-322-3676-4_16

Kulkarni KP, Kim M, Shannon JG, Lee JD (2016) Identification of quantitative trait loci controlling soybean seed weight in recombinant inbred lines derived from PI 483463 (*Glycine soja*) × 'Hutcheson'(*G. max*). Plant Breed 135(5):614–620. https://doi.org/10.1111/pbr.12407

Kulkarni KP, Asekova S, Lee DH, Bilyeu K, Song JT, Lee JD (2017) Mapping QTLs for 100-seed weight in an interspecific soybean cross of Williams 82 (*Glycine max*) and PI 366121 (*Glycine soja*). Crop Pasture Sci 68(2):148–155. https://doi.org/10.1071/CP16246

Kulkarni KP, Tayade R, Asekova S, Song JT, Shannon JG, Lee JD (2018) Harnessing the potential of forage legumes, alfalfa, soybean, and cowpea for sustainable agriculture and global food security. Front Plant Sci 9:1314. https://doi.org/10.3389/fpls.2018.01314

Lee YG, Jeong NH, Kim JH, Lee KH, Kim KH, Pirani A et al (2015) Development, validation and genetic analysis of a large soybean SNP genotyping array. Plant J 81:625–636. https://doi.org/10.1111/tpj.12755

Li H, Ye G, Wang J (2007) A modified algorithm for the improvement of composite interval mapping. Genetics

175(1):361–374. https://doi.org/10.1534/genetics.106.066811

Li YH, Zhao SC, Ma JX, Li D, Yan L, Li J et al (2013) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. BMC Genom 14:579. https://doi.org/10.1186/1471-2164-14-579

Liu HL (1949) Inheritance of defective seed coat in soybean. J Hered 40:317–322. https://doi.org/10.1093/oxfordjournals.jhered.a105970

Masuda T, Goldsmith PD (2009) World soybean demand: an elasticity analysis and long-term projections. Agricultural and Applied Economics Association: conference paper. 2009 annual meeting July 26–28. https://doi.org/10.22004/ag.econ.49490

Mao T, Li J, Wen Z, Wu T, Wu C, Sun S, Jiang B, Hou W, Li W, Song Q, Wang D, Han T (2017) Association mapping of loci controlling genetic and environmental interaction of soybean flowering time under various photo-thermal conditions. BMC Genom 18:415. https://doi.org/10.1186/s12864-017-3778-3

Markovic O, Obendorf RL (2008) Soybean seed pectinesterase. Seed Sci Res 8(4):455–461. https://doi.org/10.1017/S0960258500004426

Meng L, Li H, Zhang L, Wang J (2015) QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J 3(3):269–283. https://doi.org/10.1016/j.cj.2015.01.001

Molnar SJ, Rai S, Charette M, Cober ER (2003) Smple sequence repeat (SSR) markers linked to *E1, E3, E4*, and *E7* maturity genes in soybean. Genome 46:1024–1036. https://doi.org/10.1139/g03-079

Oyoo ME, Benitez ER, Matsumura H, Takahashi R (2010) QTL analysis of seed coat cracking in soybean. Crop sci 50:1230–1235. https://doi.org/10.2135/cropsci2009.08.0440

Saruta M, Takada Y, Yamashita K, Sayama T, Komatsu K (2019) A QTL associated with high seed coat cracking rate of a leading Japanese soybean variety. Breed Sci 69:665–671. https://doi.org/10.1270/jsbbs.19094

Shibles RM, Harper JE, Wilson RF, Shoemaker RC (2004) Soybeans : improvement, production, and uses, 3rd edn, vol 16. AGRONOMY, Madison, pp82–89

Takahashi R (1997) Association of soybean genes *I* and *T* with low-temperature induced seed coat deterioration. Crop Sci 37:1755–1759. https://doi.org/10.2135/cropsci1997.0011183X003700060014x

Takahashi R, Abe J (1999) Soybean maturity genes associated with seed coat pigmentation and cracking in response to low temperatures. Crop Sci 39:1657–1662. https://doi.org/10.2135/cropsci1999.3961657x

Toda K, Yang D, Yang D, Watanabe S, Harada K, Takahashi R (2002) A single-base deletion in soybean flavonoid 3'-hydroxylase gene is associated with gray pubescence color. Plant Mol Biol 50:187–196. https://doi.org/10.1023/A:1016087221334

Toker C (2004) Estimates of broad-sense heritability for seed yield and yield criteria in faba bean (*Vicia faba* L.). Hereditas 140(3):222–225. https://doi.org/10.1111/j.1601-5223.2004.01780.x

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93(1):77–78

Wang J, Li H, Zhang L, Meng L (2016) User's manual of QTL IciMapping ver. 4.1. The quantitative genetics group, Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081

Watanabe S, Xia Z, Hideshima R, Tsubokura Y, Sato S, Yamanaka N et al (2011) A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* Gene is involved in soybean maturity and flowering. Genet Soc America 188:395–407. https://doi.org/10.1534/genetics.110.125062

Yaklich RW, Barla-Szabo G (1993) Seed coat cracking in soybean. Crop sci 33:1016–1019. https://doi.org/10.2135/cropsci1993.0011183X003300050028x

Yang D, Nakamura T, Ohtsubo N, Takahashi K, Abe J, Takahashi R (2002) Seed coat cracking in soybean isolines for pubescence color and maturity. Crop Sci 42:71–75. https://doi.org/10.2135/cropsci2002.7100