Physiol Mol Biol Plants (November 2020) 26(11):2291–2299
https://doi.org/10.1007/s12298-020-00906-y

RESEARCH ARTICLE

Genetics and mapping of seed coat impermeability in soybean using inter-specific populations

Subhash Chandra[1,2] · Yashpal Taak[1] · Darasing Ramsing Rathod[3] · Raju Ratan Yadav[1] · Shatakshi Poonia[1] · V. Sreenivasa[4] · Akshay Talukdar[1]

Received: 3 June 2020 / Revised: 14 September 2020 / Accepted: 23 October 2020 / Published online: 6 November 2020

Abstract Seed coat impermeability (SCI) in soybean is associated with seed viability under storage and quality of processed products. Understanding genetics and identification of linked molecular markers would facilitate need-based utilization of seed coat impermeability. Two impermeable wild type (*G. soja* Sieb. and Zucc.) accessions viz. PI 424079 and PI 136620 were crossed with a permeable cultivated (*G. max*) variety JS335 to generate the mapping populations. Genetic analysis of the $F_{1:2}$ and $F_{2:3}$ seeds of the crosses indicated that SCI is controlled by a single gene/major QTL, and impermeability is dominant over permeability. Presence of seeds with intermediate permeability indicated role of some minor genes/QTLs. A set of 204 inter-specific recombinant inbred line (RILs) (F_7) was used to map SCI with 207 SSR markers. Phenotyping through rapid imbibition approach (seed imbibition for 6 h), seven QTLs were mapped on chromosomes (Chrs.) 2, 5, 12, 13 and 16 in the seeds stored for 1–3 years, while through slow imbibition method (seed imbibition for 7 days), five QTLs were mapped on Chrs. 2, 9, 10 and 20. Phenotypic variation explained (PVE) by the QTLs ranged from 5.96 to 39.67%. Two major and stable QTLs viz., qScI-h2-1 and qScI-h2-2 that mapped in tandem on Chr.2 jointly explained 43.09–62.92% of the variations in impermeability. Seven minor QTLs identified here were novel and two (qScI-h5, and qScI-h16) were consistent. It is the first report of mapping impermeability using two imbibition approaches together in 200 plus inter-specific RILs in soybean. The study will pave the way for developing genotypes with restricted permeability, enhanced seed viability, and improved seeds quality.

Keywords Soybean · Seed coat impermeability · Inheritance · QTL · RIL · SSR markers

Introduction

Cultivated soybean (*Glycine max* L. Merrill) is believed to have domesticated from its wild annual counterpart *Glycine soja* (Sieb. and Zucc.) in East Asia (Hymowitz and Singh 1987). Both the species differ for a multitude of traits including seed coat permeability; the seeds of *G. soja* are impermeable while the seeds of the cultivated types are permeable (Mullin and Xu 2000; Chandra et al. 2017). The 'impermeability' or 'hard-seededness', which are used interchangeably in literature indicates inhibition of imbibitions (Liu et al. 2007; Jang et al. 2015). It has been associated with some positive impacts to the seeds such as tolerance to fungal activity (Starzinger and West 1982) and flooding during germination (Hou and Thseng 1991). It also enhances seed longevity under delayed-harvest field conditions (Hartwig and Potts 1987) and humid environments (Potts et al. 1978) by restricting rapid imbibition of water. Contrarily, it is an undesirable trait for the soybean processing industries because it reduces the marketing quality of the processed products (Ma et al. 2004).

Experiments conducted to understand the genetics of seed coat impermeability have reported diverse results.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12298-020-00906-y) contains supplementary material, which is available to authorized users.

✉ Akshay Talukdar
akshay.talukdar1@gmail.com

[1] Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

[2] ICAR-Indian Institute of Soybean Research, Indore, India

[3] P.D. Krishi Vishwavidyalaya, Akola, India

[4] ICAR-Sugarcane Breeding Institute, Coimbatore, India

Kilen and Hartwig (1978) suggested that three major genes control the permeable—impermeable responses with involvement of some minor genes. Watanabe et al. (2004) and Sakamato et al. (2004) reported involvement of three and four QTLs, respectively while Liu et al. (2007) and Zhang et al. (2008) reported involvement of two QTLs. Many genomic regions on a number of chromosomes viz., 2, 3, 6, 8, 14, 18, 19 and 20 have also been reported to influence seed coat impermeability in soybean (Keim et al. 1990; Sakamoto et al. 2004; Liu et al. 2007; Singh et al. 2008; Ai et al. 2018; Chen et al. 2019). Kebede et al. (2014) and Sun et al. (2015) indicated monogenic control of this trait. While studying molecular basis of hard seededness, Jang et al. (2015) observed that a single nucleotide substitution i.e. A with G at position no. 863bp in the nucleotide sequence of *qHS1*altered an impermeable seed to permeable one. In a similar study, Sun et al. (2015) observed that C > T mutation in the locus *GmHs1-1* coupled with some minor QTLs can transform the hard seeds to soft seeds.

Incorporation of some degree of seed coat impermeability is a potential approach of retarding deteriorative process in the field and storage without compromising quality of seeds and soy-based products (Potts et al. 1978; Kilen and Hartwig 1978). To achieve this goal, knowing genetics of SCI and molecular markers linked to it would be critical. Previous attempts have mapped SCI into various genomic regions across the soybean genome. Such studies varied not only in the mapping population used but also in the phenotypic scale used in categorizing the seeds as permeable and impermeable. Sun et al. (2015) considered seeds up to 20% imbibition as permeable while Kebede et al. (2014) kept the cut-off mark at 4%. Similarly, imbibition period of the seeds also found to vary from 2–6 h. (rapid approach) to 7 days (slow approach) (Sakamoto et al. 2004; Liu et al. 2007; Kebede et al. 2014; Sun et al. 2015; Chen et al. 2019). In the present study, attempts were made to understand inheritance of SCI in the *G. soja* accessions and to map the genomic region(s) for SCI using SSR markers. Unlike other reports, this study used a 200 plus inter-specific recombinant inbred line (RIL) population and adopted both rapid and slow approaches of imbibition in seeds stored for 3 years of ambient storage for mapping purpose.

Material and methods

Plant material

Two *G. soja* accessions viz., PI 424079 and PI 366120 were crossed in pair with a *G. max* variety JS 335. Seeds of the *G. soja* accessions were small and impermeable while *G. max* seeds were permeable (Chandra et al. 2017). For effective hybridization, *G. soja* genotypes were used as female. Since, seed coat is a maternal tissue, therefore seed coat phenotype of $F_{1:2}$ and $F_{2:3}$ seeds were used to deduce the genotypes of the F_1 and F_2 plants, respectively. A total of 124 and 93 plants constituted the experimental population of PI 424079 × JS 335 and PI 366120 × JS 335, respectively.

For mapping purpose, an inter-specific RIL developed by crossing DC 2008-1, a *G. soja* accession (impermeable seed coat) with DS 9712, a *G. max* variety (permeable seed coat) of North India was used. The 204 RILs in F_7 were used to map the genomic regions controlling seed coat impermeability. Freshly harvested seeds of 204 RILs and their parents were air dried to about 10% moisture contents and packed in cloth bags and stored under ambient conditions of Delhi for different periods (1, 2 and 3 years) with an average relative humidity of 65 ± 5% and temperature of 25 ± 2 °C as described by Kumar et al. (2019).The results obtained on the RILs were validated in a F_2 population derived from a cross involving seed-coat impermeable genotype PI 424079 with permeable genotype JS 335.

Seed coat impermeability assay and segregation analysis

To test seed coat impermeability, seeds of the parental genotypes along with the $F_{1:2}$ and $F_{2:3}$ seeds of both the crosses (PI 424079 × JS 335 and PI 366120 × JS 335) were subjected to rapid imbibition test as per Jang et al. (2015), where 20 seeds were soaked in 100 ml distilled water for 6 h at room temperature. Seed coat impermeability was recorded as the percentage of hard seeds that had not imbibed water after soaking for 6 h and was designated as ScI-h (Seed coat impermeability based on hours). A genotype with ≤20% imbibed seeds was considered as impermeable while that with ≥80% was recorded as permeable (Sun et al. 2015). To maximize the likelihood of imbibition being the result of seed-coat permeability rather than seed-coat cracking or damage, only seeds with intact seed coats were chosen and used. Accordingly, seeds of the F_2 plants were grouped as impermeable or hard-seeded and permeable or soft-seeded. Gene models viz., one gene model, two gene model and three gene model were tested to estimate the number of gene(s) governing seed coat impermeability. Chi-square (χ^2) test was applied to the observed segregation data in F_2 populations to test the goodness of fit of various probable genetic ratios.

Phenotyping of RILs and statistical analysis

Seed coat impermeability of the 204 RILs along with the parental genotypes was estimated through rapid imbibition test (50 seeds in three replications) as described above (Jang et al. 2015). Additionally, the seeds of the RILs were subjected to slow imbibition test. For it, 50 seeds of each genotype in three replications were placed in water-moistened germination papers, rolled up and positioned vertically in a germination chamber at 25 °C. The germination papers were maintained at near 100% humidity for 7 days. At the end of the 7th day, seeds were checked for imbibition. The seeds that did not imbibe water were counted as "impermeable or hard seed," and that imbibed water were counted as "permeable or soft seeds" as per Kebede et al. (2014). The seed coat impermeability (%) data of slow approach were collected through Complete Randomized Design (CRD) and designated as ScI-d (Seed coat impermeability based on days). The data of seed coat impermeability (%), seed coat colour and 100 seed weight (g) were recorded for the seeds stored for 1–3 years. Mean value of both seed coat impermeability (ScI-h and ScI-d) for three different environments (3 storage periods) were used for analyzing and mapping QTL for the SCI.

Genetic polymorphism survey, QTL analysis and validation

For genetic polymorphism survey between the parental genotypes of the RILs i.e. DC 2008-1 and DS 9712, 400 simple sequence repeat (SSR) markers were picked up from the consensus genetic map of soybean (approx. 8–12 markers per arm of a chromosome). Genomic DNA was isolated from the young leaves of the parental genotypes and the RILs following CTAB (cetyltrimethyl ammonium bromide) procedure (Saghai-Maroof et al. 1984). Purified DNA was subjected to PCR amplification by SSRs as per protocol described by Yashpal et al. (2019). Amplified products were resolved on 3% high resolution agarose (Metaphore™ Lonza, Switzerland) gels stained with Ethidium Bromide in 1 × TAE buffer at 80 V for 2.5–3.0 h as per Kumawat et al. (2015), Yashpal et al. (2019), and analysed in gel documentation system (AlphaImager-Protien Simple, USA). Out of 400 SSR markers, 218 appeared to be polymorphic between the parental genotypes, which were subsequently used for genotyping the RILs. The genotypic data were checked through χ^2 test for segregation distortion. Eleven SSR markers were rejected due to complex banding patterns and distorted segregations, and the rest (207 markers) were used to construct the linkage map using the QTL IciMapping software (Meng et al. 2015).

The QTL for SCI analysis in each year of testing was performed using both the phenotyping systems i.e. ScI-h and ScI-d using inclusive composite interval mapping (ICIM) approach (Wang 2009) using the software QTL IciMapping (Meng et al. 2015). The scanning interval was set to 1 cM, and the probability of stepwise regression for markers entering the model was set to 0.001. A manual input of 3.0 was set for the threshold LOD score to declare a significant QTL. Mean values of the phenotypic data were used for mapping and estimating the phenotypic variation explained (PVE) and the additive effect of the QTLs as per Wang (2009). A QTL was considered stable if it was detected in at least 2 of the 3 years of testing. QTL nomenclature was followed as per McCouch et al. (1997). The QTLs expressing over the years of testing were used to validate in a set of 124 F_2 plants of the cross PI 424079 (impermeable) × JS 335 (permeable). The mean seed coat impermeability data (ScI-h) of the F_2 plants along with the identified markers were subjected to map QTL following single marker analysis (SMA) approach using QTL IciMapping software.

Results

Segregation analysis in F_2 plants

The seed coat impermeability index in the seeds of JS 335 were 0 i.e. the seed coats were highly permeable while that in the seeds of PI 424079 and PI 366120 was 100% i.e. highly impermeable. All the F_1 seeds derived from both crosses i.e. PI 424079 × JS 335 and PI 366120 × JS 335 were impermeable. The seed coat impermeability in the 136 $F_{1:2}$ seeds of the cross PI 424079 × JS 335 was 84–89% (av. 87%) while that in 112 $F_{1:2}$ seeds of the cross PI 366120 × JS 335 was 70–84% (av. 82%). Thus, the seed coat of the $F_{1:2}$ seeds in both crosses were impermeable in nature indicating impermeability to be dominant over permeability.

Seed coat impermeability in the $F_{2:3}$ seeds of both the crosses ranged from 0 to 100%. The frequency distributions of the impermeable and permeable seeds exhibited a bi-modal pattern with more numbers in the two extreme classes viz. permeable and impermeable with a few in the intermediates (Fig. 1). In the cross PI 424079 × JS 335, the number of F_2 plants having impermeable and permeable seeds were 63 and 24, respectively. Similarly, in the cross PI 366120 × JS 335, the number of F_2 plants with impermeable and permeable seeds were 45 and 19, respectively.

The test for goodness of fit for both the crosses through χ^2 (chi-square) test showed that the data fits significantly to the 3 impermeable:1 permeable ratio (Table 1) indicating

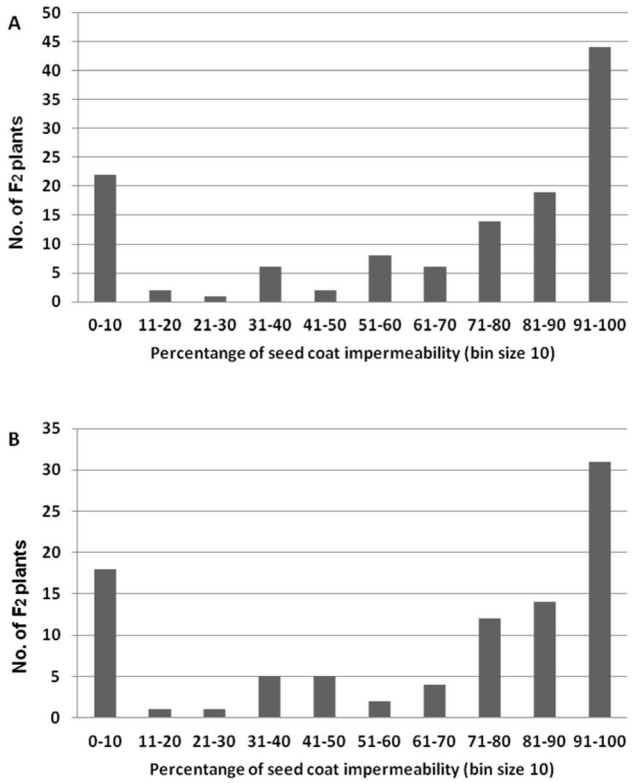

Fig. 1 Frequency distributions for seed coat impermeability in F_2 individuals of cross **a** PI 424079 × JS 335 and **b** PI 366,120 × JS 335

involvement of a single dominant gene in controlling seed coat impermeability or hard seededness in the wild type soybean (*G. soja*) genotypes PI 424079 and PI 366120. However, in both the crosses, 36 and 32 F_2 plants with intermediate seed coat permeability were also observed, which were eliminated from the analysis. Emergence of seed coat with intermediate permeability indicated that besides the single major gene, some other minor gene(s) might be involved in controlling the seed coat impermeability in the soybean seeds.

Phenotyping of recombinant inbred lines

The RIL was developed from a cross involving DS 9712 (*G. max*) and DC 2008-1 (*G. soja*) genotypes. The DS 9712 was erect plant with determinate growth habit. It has large seeds (mean 100-seed weight 9.12 g) with yellow and permeable seed coat. The wild type accession DC 2008-1 was a climber with indeterminate growth habit. The seeds of DC 2008-1 were black, exceedingly small (mean 100-seed weight 0.56 g) with impermeable seed coat. The impermeability index of DS 9712 was 0.9% and 0.2% for ScI-h and ScI-d respectively, while that of DC 2008-1 was 100% and 74% for ScI-h and ScI-d, respectively.

The ANOVA of the 204 RILs for ScI-h and ScI-d indicated that there were significant variations among the RILs for impermeability, different storage periods and interactions between the genotypes and storage periods (Supplementary Table 1). For the rapid test, the seed coat impermeability index (ScI-h) of the RILs ranged from 0 to 100%. The mean SCI of the seeds were 46.17%, 54.42% and 60.75% after ambient storage for 3, 2 and 1 year, respectively. Out of the total RILs, 45 RILs were classified as permeable and 60 as impermeable following the scale given by Sun et al. (2015). Similarly, based on slow-test, the SCI index (ScI-d) ranged from 0–90%, 0–72% and 0–68% with mean value of 12.5%, 15.28%, and 21.15% in the seeds stored for 3 years, 2 years and 1 year, respectively. Following the scale of Sun et al. (2015), 139 RILs were categorized as permeable but none could be classified as impermeable. In the slow imbibition test, a few RILs appeared to have intermediate permeability. When the data of intermediate permeability were used, the frequency distribution curve of SCI indices turned from bi-modal to uni-modal (Fig. 2). It was also observed that the SCI index value decreased with the period of seed storage i.e. with the increase in storage period, the seed coats became increasingly permeable.

In correlation analysis, it was observed that the SCI index ScI-h was positively and significantly correlated (r = 0.721**) with ScI-d. Both the indices i.e. ScI-h and ScI-d were positively correlated (r = 0.156*) with seed size

Table 1 Chi-square (χ^2) tests for goodness of fit for seed coat impermeability

Genotype	Range of impermeable seed (%)	Impermeable*F_2 plants (No.)		Permeable*F_2 plants (No.)		χ^2 (3:1)	*P*-value
		Observed	Expected	Observed	Expected		
PI 424079 × JS 335							
F_2	0–100	63	65.25	24	21.75	0.31	0.58
PI 366120 × JS 335							
F_2	0–100	45	48	19	16	0.75	0.39

*Seed coat impermeability with ≤ 20% is classified as permeable and ≥ 80% as impermeable genotype as per Sun et al. (2015)

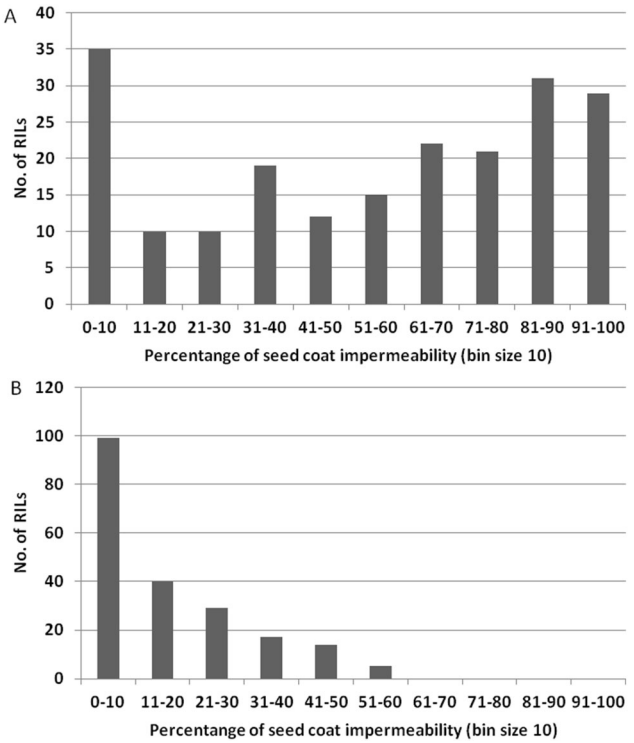

Fig. 2 Frequency distributions for seed coat impermeability in RILs based on mean value over the years; **a** based on rapid imbibition test (ScI-h); **b** based on slow imbibition test (ScI-d)

but not with seed coat color (Supplementary Figure 1). Based on ScI-h, the RILs with permeable seed coat included 18 lines with green seeds, 14 with black, 8 yellow and 5 brown seeds. Similarly, the RILs with impermeable seed coat included 37 black, 16 green, 4 yellow and 3 brown seeded lines. Although the colour of the seeds found to have no specific association with permeability of the seed coat; however, the number of impermeable seeds with black seed coat was more in numbers than others.

Mapping the QTLs for seed coat impermeability

Out of the 400 SSR markers used in polymorphic study, 218 appeared to be polymorphic between the two genotypes indicating the level of polymorphism between them to be 54.56%. The distribution of the polymorphic markers across the genome, however, were not uniform, it varied from one chromosome to the other. Highest level of polymorphism (72.22%) was observed on Chr.19, while the least (45.45%) was observed on Chr.6. The linkage map had a length of 1851.55 cM with an average marker distance of 8.94 cM. Order of the markers on the map was the same as in the reference map with occasional variation in the inter-marker map distance in the map. Genomic regions affecting seed coat impermeability were mapped based on both, rapid imbibition (ScI-h) and slow imbibition (ScI-d) data. The phenotypic and genetic data indicated that besides major gene, some minor genes were also involved in controlling the trait. Therefore, the mapping of the genomic regions for seed coat impermeability was processed through QTL mapping approach.

Based on rapid imbibition approach, seven QTLs viz. *qScI-h 2-1*, *qScI-h 2-2, qScI-h 5-1, qScI*-h 5-*2, qScI-h 12, qScI-h 13* and *qScI-h 16* were mapped on five chromosome viz., chr.2, chr.5, chr.12, chr.13 and chr.16, respectively. Among these seven QTLs, only one i.e. *qScI-h 2-2* was mapped consistently over the 3 years of testing, whereas, *qScI-h 2-1* and *qScI-h 16* were mapped over 2 years of testing, and rest four QTLs were mapped on a particular year only (Table 2). LOD score of the QTL ranged from 3.01 (*qScI-h 13*) to 17.74 (*qScI-h 2-2*) (Supplementary Figure 2). The two QTLs mapped on Chr.2 i.e. *qScI-h 2-1* and *qScI-h 2-2* were positioned in tandem and were consistent over the 3 years of testing. Further, both the QTLs were flanked by one common marker Satt274. Taken together, these two QTLs explained 43.09–62.92% of the phenotypic variations of seed coat impermeability. Therefore, it can be considered that genomic region covered by the markers Satt703-Satt274-Satt202 harbors a major QTL (or a gene) that control seed coat impermeability in soybean (Fig. 3). Rest of the QTLs appeared to be minor having contribution to total phenotypic variance less than 10%. All but one QTL had positive additive effects indicating their origin to the parent having impermeable seed coat (Table 2). Of the fourteen markers linked to seven QTLs, the minimum distance (0.47 cM) between marker and QTL was observed between Sct_046 and qScI-h 16.

Based on the slow imbibition test, 5 QTLs were mapped on four chromosomes. The LOD of the QTLs mapped ranged from 3.77 (qScI-d 10) to 11.00 (qScI-d 2-2) and the phenotypic variation explained (PVE) by it ranged from 7.70% (*qScI-d 10*) to 26.64% (*qScI-d 2-1*). Out of these five QTLs mapped, two overlapped the genomic region of the QTL *qScI-h 2-1* and *qScI-h 2-2*.

The QTL *qScI-d 2-2* flanked by the markers Satt274-Sat_202 was mapped consistently over the 3 years at map position 117.23–118.23 cM. It explained about 19.80–23.30% of the phenotypic variations during the 3 years of mapping (Table 2). The proximity between QTL and the marker ranged from 0.45 cM (*qScI-d 10*-Sat_291) to 11.48 cM (*qScI-d 2-1*-Satt703).

For validating the QTLs mapped, 18 SSR markers linked to all QTLs were selected and used to genotype 124 F_2 individuals of the cross PI 424079 × JS 335. Out of these 12 QTLs, three QTLs viz., *qScI-h 2-1*, *qScI-h 2-2* and *qScI-d 2-2* were validated in the F_2 mapping population. The PVE for the validated QTLs ranged from 10.12 to 18.69%. Similarly, the SSR markers viz., Sat_202, Satt703,

Table 2 QTL(s) for seed coat impermeability (SCI) identified over the storage years

ScI Index	Seed storage period(yr.)	Chr. No.	QTL	Map Position (cM)	Left Marker	Right marker	LOD	PVE (%)	Add
ScI-h	III	2	*qScI-h 2-1*	109.23	Satt703	Satt274	6.32	16.49	14.92
		2	*qScI-h 2-2*	117.23	Satt274	Sat_202	11.83	26.66	19.03
		5	*qScI-h 5-1*	70.56	Satt619	Satt545	3.42	5.96	9.00
		16	*qScI-h 16*	38.94	Sct_046	Satt686	3.72	6.74	9.55
	II	2	*qScI-h 2-1*	109.23	Satt703	Satt274	8.78	23.25	18.32
		2	*qScI-h 2-2*	118.23	Satt274	Sat_202	17.74	39.67	24.05
		5	*qScI-h 5-2*	56.56	Sat_356	Satt619	3.04	7.86	10.64
		16	*qScI-h 16*	39.94	Sct_046	Satt686	4.36	6.70	9.84
	I	2	*qScI-h 2-2*	117.23	Satt274	Sat_202	13.64	29.38	19.89
		12	*qScI-h 12*	91.59	Satt181	Satt434	3.85	6.22	-9.15
		13	*qScI-h 13*	11.79	Satt325	Satt252	3.01	6.71	9.49
ScI-d	III	2	*qScI-d 2-1*	110.23	Satt703	Satt274	9.31	26.64	8.25
		2	*qScI-d 2-2*	117.23	Satt274	Sat_202	9.15	23.30	7.74
	II	2	*qScI-d 2-2*	118.23	Satt274	Sat_202	11.00	23.24	8.84
		9	*qScI-d 9*	103.85	Sat_352	Satt196	4.84	8.78	5.38
		20	*qScI-d 20*	88.33	Satt162	Sat_420	5.08	10.33	5.83
	I	2	*qScI-d 2-2*	118.23	Satt274	Sat_202	8.03	19.80	10.60
		10	*qScI-d 10*	51.44	Satt347	Sat_291	3.77	7.70	6.56

ScI-h: Seed coat impermeability after 6 h of imbibition, ScI-d: Seed coat impermeability after 7 days of imbibition, LOD: LOD score, PVE (%): Phenotypic variation explained by QTL, Add.: Estimated additive effect of the marker

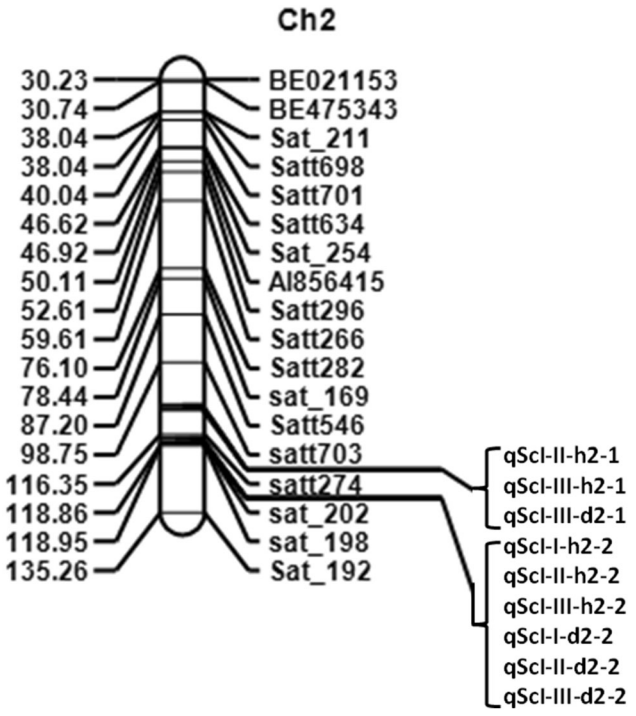

Fig. 3 Major and consistent QTLs identified for ScI-h and ScI-d (I, II, III in QTLs name represent number of years of storage)

Satt274 from the Chr.2 exhibited confirmed association with seed coat impermeability index.

Discussion

Inheritance of seed coat impermeability

Seed coat impermeability is a domestication related trait. It favored the wild type genotypes to overcome the unfavorable environmental conditions and continue the generations with enhanced viability. However, in the present-day context, it is not a highly desirable trait as it delays germination considerably and makes food processing cumbersome. During evolution and domestication, the impermeable seed coat became permeable but not without penalty; the permeable seeds tend to lose viability early. Therefore, it is imperative to develop highly viable seeds with permeable seed coat (Kumar et al. 2019). In an attempt to understand the genetic control of seed coat impermeability, it was found in two F_2 populations that a major gene controls this trait in two wild type (*G. soja*) genotypes viz., PI 424079 and PI 366120. It was also apparent from the F_1 data that the impermeability was dominant over permeability. Kebede et al. (2014) and Sun et al. (2015) also reported the monogenetic control of the seed coat impermeability in soybean. Contrarily, Kilen and

Hartwig (1978) and Keim et al. (1990) reported involvement of three and five genes, respectively, in expression of hard seededness in soybean. Such variations in the result emanated primarily because of the variation in the genotypes used for the experimentation and the classification adopted for phenotyping. Surprisingly, no phenotypic scale has yet been accepted universally for categorizing the seeds as permeable or impermeable. Sun et al. (2015) considered seeds with 0–20% imbibition as permeable while Kebede et al. (2014) kept the cut-off of imbibition at 4% only. Similarly, period of soaking in water for imbibition also found to vary from report to report. Some researchers soaked the seeds for 2–6 h (Sakamoto et al. 2004; Watanabe et al. 2004; Liu et al. 2007; Jang et al. 2015; Sun et al. 2015, Chen et al. 2019), while others soaked it for 7 days (Keim et al. 1990; Kebede et al. 2014). Such variations in assessment of the phenotype led to the differences in the ratio of impermeable to permeable seeds and eventual variation in the genetics of impermeability.

In this study, the monogenetic control of impermeability was obtained when the data were classified in to two extreme classes i.e. impermeable (80% or more impermeable seeds) and permeable (20% or less impermeable seeds). However, a non-negligible number of seeds appeared as semi-impermeable or intermediate permeable (>20 to <80%). To reduce the number of seeds in the intermediate classes, we extended the period of imbibition from 2 to 6 h as per Chandra et al. (2017); however, a few seeds remained as semi-impermeable. Inclusion of these to any classes led the data to no specific ratio, nor satisfied the polygenic control theory. Therefore, the seeds with intermediate phenotype were excluded from the classification as per Sun et al. (2015) and the monogenic control of the trait was obtained. The seeds with semi-impermeable seed coat appear because of the environment or involvement of some minor genes. Sudden change in environmental factors such as temperature leads to formation of the hard seeds (Egli et al. 2005). However, the present experiment was conducted in controlled environmental condition and hence role of the adverse environment in making the seed coat impermeable can be ruled out. The seeds in the $F_{1:2}$ generation was impermeable rather than intermediate. It indicated the dominant nature of the gene rather than additive. Similarly, the $F_{2:3}$ data showed a bimodal distribution rather than continuous variation. Such observations clearly indicated that the seed coat impermeability is under control of a major gene; however, involvement of some minor genes could not be ruled out, because many seeds with intermediate impermeability also appeared. Therefore, it would be convenient to consider the 'major gene' as a 'major QTL' and the other genes as 'minor QTLs'. Degree of impermeability in the seed coat would vary with the number of minor QTLs involved along with the major QTL. Absence of the major QTL along with the minors would make a seed coat permeable and vice-versa.

Mapping of seed coat impermeability

Seed coat impermeability is a seed trait. One would need to wait for one complete generation i.e. from seed to seed to see the changes in the seed trait. Therefore, it would be quite useful and time saving, if marker(s) linked to seed coat impermeability becomes available. Reports are available about the mapping and identification of markers linked to seed coat impermeability; however, inconsistency in the result is quite common. It might have happened for several reasons including the genotypes, mapping population and the marker system used in mapping. Variations have also been observed in the phenotypic approach adopted. The rapid imbibition approach i.e. soaking the seeds for 2–6 h was adopted for phenotyping by Sakamoto et al. (2004), Watanabe et al. (2004), Liu et al. (2007), Jang et al. (2015) and Sun et al. (2015), while slow imbibition approach i.e. testing the seeds in paper towels for about 7 days was adopted by some others (Keim et al. 1990; Kebede et al. 2014). However, we did not find any report where both the phenotypic approaches were used together and compared the result. In this study, the seeds were phenotyped using both the methods of imbibition and mapping were applied accordingly. It was observed that 7 QTLs were mapped when phenotyping was done through the rapid imbibition approach, while the number of QTLs came down to 5 in the other approach. Some minor QTLs that were mapped in the rapid imbibition approach missed on the other approach. The PVE (%) value of the QTLs also changed with the change of approach. It was due to phenotyping variations for both the testing approaches i.e. ScI-h and ScI-d. However, position of the mapped QTLs remained nearly unchanged. Since the results are comparable, therefore, anyone of these two approaches can be used for mapping; however, in the absence of constraints one would better go for slow imbibition approach to test permeability of the seed coat and map the QTLs.

In this study, 12 QTLs for seed coat impermeability were mapped using rapid and slow process of imbibition over 3 years of testing. The mapped QTLs found to spread over 8 different chromosomes. While comparing this result with the previous reports, it was found that QTLs mapped on chr.2 and chr.20 were also reported earlier by several workers (Keim et al. 1990; Sakamoto et al. 2004; Watanabe et al. 2004; Liu et al. 2007; Kebede et al. 2014). It thus confirms correctness of the current result. The genomic region harboring the QTL for SCI on chr.2 is represented by the SSR markers Satt703, Satt274 and Sat_202 at map position 98.75, 116.34 and 118.86 cM, respectively (www.soybase.org). Previous reports that mapped QTLs for SCI

in this region of chromosome 2 varied in the PVE (%) which accounted for 13% (Keim et al. 1990), 10.6% (Watanabe et al. 2004), 38.5% (Sakamoto et al. 2004), 47.8% (Liu et al. 2007) and 63–92% (Kebede et al. 2014). Sun et al. (2015) and Jang et al. (2015) detected one SNP (A > G) in this locus that explained 60–70% variation for this trait. In the present study, two QTLs (*qScI-h 2-1* and *qScI-h 2-2*) were mapped in tandem on this region of chr.2. Taken together both the QTLs explained 43–63% of the phenotypic variations. Based on the PVE (%), it was considered that the QTL mapped on the chr.2 is a major QTL. Since contribution of this QTL is quite high, hence it also appears as a major gene. Thus, it was confirmed that the seed coat impermeability is controlled by a major QTL along with some minor QTLs.

Detection of novel minor QTLs

In this study, seven minor QTLs were mapped on seven different chromosomes. All these seven QTLs were different from those reported earlier elsewhere for the trait under consideration and hence could be considered as novel QTLs. Two of the minor QTLs viz., *qScI-h 5*, and *qScI-h 16* were mapped consistently over 2 years of testing explaining 5–6% of the variations in seed coat impermeability. Similarly, the QTLs mapped on the chr.2 were also appeared to be stable with higher contribution towards seed coat impermeability. Stably expressing QTL ought to be preferred over an unstable QTL even if its effect is moderate (Liu et al. 2007). Therefore, the two major QTLs mapped on chr.2 and the two minor QTLs mapped on chr.5 and chr.16 holds promise for future use. The genomic region represented by the QTL *qScI-h 16* also reported to harbor QTL for leaflet width, phytophthora resistance, seed tocopherol etc. Similarly, the genomic regions near to *qScI-h 5-1* co-holds QTLs for whitefly resistance, seed genitein and seed daidzein (www.soybase.org). Therefore, deployment of such QTLs for partial impermeability is expected to benefit the genotype with other desirable traits as well.

Validation

Validation of identified QTLs in a set of unrelated germplasm or mapping population reflects true value of the QTLs mapped and its applicability. In the present study, applicability of the SSR markers linked to the QTL for seed coat permeability were validated by single marker analysis in an unrelated F_2 population developed for the same trait. It was found that three linked markers could rightly identify the target QTL and its phenotypic variations ranged from 10.12% to 18.96%. It thus proved applicability of these markers in deployment of this major QTL through MAS. Similar approach of QTL validation was also adopted by Kebede et al. (2014). The minor QTLs mapped in the study could not be confirmed in the validation population. Such variations arise due to change in mapping population (Radhika et al. 2007). It may be due to the background effect i.e. presence of alleles other than the one mapped in original mapping population (Palomeque et al. 2010).

The QTL mapped and the linked markers (Sat_202, Satt703, Satt274, Satt686, Satt619 and sct_046) identified in this study would be useful in making any genotype semi-impermeable and highly viable. Makers would also be useful for identification of seed coat permeable segregants at seedling stage. Thus the information generated in this study would open vistas for developing soybean genotypes with high viability without compromising quality of the processed products.

Acknowledgements First author is grateful to the UGC, New Delhi for Junior Research Fellowship during the doctoral study and authors are grateful to the National Agricultural Science Fund, ICAR, New Delhi for funding the program.

Authors' contribution Conceptualization of research (AT); Designing of the experiments (AT, SC); Contribution of experimental materials (AT, YT); Execution of field/lab experiments and data collection (SV, SP, RRY, SC, DRR); Analysis of data and interpretation (SC,YT, AT); Preparation of the manuscript (SC, AT).

References

Ai LJ, Chen Q, Yang CY, Yan L, Wang FM, Ge RC, Zhang MC (2018) Additive and epistatic QTL mapping for soybean hard seededness. Acta Agron Sin 44(6):852–858

Chandra S, Yadav RR, Poonia S, Yashpal DR, Rathod A, Kumar SKL, Talukdar A (2017) Seed coat permeability studies in wild and cultivated species of soybean. Int J Curr Micro Appl Sci 6(7):2358–2363

Chen JJ, Liu XX, Yu LL, Lu YP, Zhang ST, Zhang HC, Guan RX, Qiu LJ (2019) QTL mapping of hard seededness in wild soybean using BSA method. Sci Agri Sin 52(13):2208–2219

Egli DB, Tekrony DM, Heitholt JJ, Rupe J (2005) Air temperature during seed filling and soybean seed germination and vigor. Crop Sci 45:1329–1335

Hartwig EE, Potts HC (1987) Development and evaluation of impermeable seed coats for preserving soybean seed quality. Crop Sci 27:506–508

Hou FF, Thseng FS (1991) Studies on the flooding tolerance of soybean seed: varietal differences. Euphytica 57:l69-173

Hymowitz T, Singh RJ (1987) Taxonomy and speciation. In: Wilcox JR (ed) Soybeans: Improvement, Production and Uses, 2nd edn. American Society of Agronomy, Madison, WI, pp. 23–48

Jang SJ, Sato M, Sato K, Jitsuyama Y, Fujino K, Mori H (2015) A single nucleotide polymorphism in an endo-1, 4-β-glucanase gene controls seed coat permeability in soybean. PLoS One 10(6):e0128527

Kebede H, Smith JR, Ray JD (2014) Identification of a single gene for seed coat impermeability in soybean PI 594619. Theor Appl Genet 127(9):1991–2003

Keim P, Diers BW, Shoemaker RC (1990) Genetic analysis of soybean hard seededness with molecular markers. Theor Appl Genet 79:465–469

Kilen TC, Hartwig EE (1978) An inheritance study of impermeable seed in soybean. Field Crop Res 1:65–70

Kumar A, Chandra S, Talukdar A, Yadav RR, Saini M, Poonia S, Lal SK (2019) Genetic studies on seed coat permeability and viability in RILs derived from an inter-specific cross of soybean [Glycine max (L.) Merrill]. Indian J Genet 79(1):48–55

Kumawat G, Singh G, Gireesh C, Shivakumar M, Arya M, Agarwal DK, Husain SM (2015) Molecular characterization and genetic diversity analysis of soybean [*Glycine max* (L.) Merr.] germplasm accessions in India. Physiol Mol Biol Plants 21:101–107

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

Ma F, Cholewa E, Mohamed T, Peterson CA, Gijzen M (2004) Cracks in the palisade cuticle of soybean seed coats correlates with their permeability to water. Ann Bot 94:213–228

McCouch SR, Cho YG, Yano M, Paul EM, Blinstrub M, Morishima H, Kinoshita T (1997) Report of QTL nomenclature. Rice Genet Newsl 14:11–13

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 3:269–283

Mullin WJ, Xu WL (2000) A study of the intervarietal differences of cotyledon and seed coat carbohydrates in soybean. Food Res Int 33:883–891

Palomeque L, Liu LJ, Li W, Hedges BR, Cober ER, Smid MP, Lukens L, 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

Potts HC, Duangpatra J, Hairston WG, Delouche JC (1978) Some influences of hardseededness on soybean seed quality. Crop Sci 18:221–224

Radhika P, Gowda SJM, Kadoo NY, Mhase LB, Jarnadagni BM, Sainani MN, Chandra S, Gupta VS (2007) Development of an integrated intraspecific map of chickpea (*Cicer arietinum* L.) using two recombinant inbred line populations. Theor Appl Genet 115:209–216

Sakamoto S, Abe J, Kanazawa A, Shimamoto Y (2004) Marker-assisted analysis for hard seededness with isozyme and simple sequence repeat loci. Breed Sci 54:133–139

Saghai-Maroof MA, Soliman KM, Jergensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci 81:8014–8018

Singh RK, Raipuria RK, Bhatia VS, Rani A, Husain SM, Satyavathi CT, Chauhan GS, Mohapatra T (2008) Identification of SSR markers associated with seed coat permeability and electrolyte leaching in soybean. Physio Mol Bio Plants 14(3):173–177

Starzinger EK, West SH (1982) An observation in the relationship of soybean seed colour to viability maintenance. Seed Sci Technol 10:301–305

Sun L, Miao Z, Cai C, Zhang D, Zhao M, Wu Y, Zhang X, Swarm SA, Zhou L, Zhang ZJ et al (2015) GmHs1-1, encoding a calcineurin-like protein, controls hard-seededness in soybean. Nature Genet 47(8):939–943

Wang J (2009) Inclusive composite interval mapping of quantitative trait genes. Acta Agron Sin 35:239–245

Watanabe S, Tajuddin T, Yamanaka N, Hayashi M, Harada K (2004) Analysis of QTLs for reproductive development and seed quality traits in soybean using recombinant inbred lines. Breed Sci 54:399–407

Yashpal DR, Rathod S, Chandra A, Kumar RRY, 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

Zhang B, Chen P, Chen CY, Wang D, Shi A, Hou A, Ishibashi T (2008) Quantitative trait loci mapping of seed hardness in soybean. Crop Sci 48:1341–1349