



Fine-mapping and identifying candidate genes conferring resistance to *Soybean mosaic virus* strain SC20 in soybean

Adhimoolam Karthikeyan^{1,2,3,4} · Kai Li^{1,2,3,4,5} · Cui Li^{1,2,3,4} · Jinlong Yin^{1,2,3,4} · Na Li^{1,2,3,4} · Yunhua Yang^{1,2,3,4} · Yingpei Song^{1,2,3,4} · Rui Ren^{1,2,3,4} · Haijian Zhi^{1,2,3,4,5} · Junyi Gai^{1,2,3,4,5}

Received: 2 July 2017 / Accepted: 4 November 2017 / Published online: 27 November 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

Key message The Mendelian gene conferring resistance to *Soybean mosaic virus* Strain SC20 in soybean was fine-mapped onto a 79-kb segment on Chr.13 where two closely linked candidate genes were identified and qRT-PCR verified.

Abstract *Soybean mosaic virus* (SMV) threatens the world soybean production, particularly in China. A country-wide SMV strain system composed of 22 strains was established in China, among which SC20 is a dominant strain in five provinces in Southern China. Resistance to SC20 was evaluated in parents, F₁, F₂ and the F_{2:7} RIL (recombinant inbred line) population derived from a cross between Qihuang-1 (resistant) and NN1138-2 (susceptible). The segregation ratio of resistant to susceptible in the populations suggested a single dominant gene involved in the resistance to SC20 in Qihuang-1. A “partial genome mapping strategy” was used to map the resistance gene on Chromosome 13. Linkage analysis between 178 RILs and genetic markers showed that the SC20-resistance gene located at 3.9 and 3.8 cM to the flanking markers BARCSOYSSR_13_1099 and BARCSOYSSR_13_1185 on Chromosome 13. Subsequently, a residual heterozygote segregating population with 346 individuals was developed by selfing four plants heterozygous at markers adjacent to the tentative SC20-resistance gene; then, the candidate region was delimited to a genomic interval of approximately 79 kb flanked by the new markers gm-ssr_13-14 and gm-indel_13-3. Among the seven annotated candidate genes in this region, two genes, *Glyma.13G194700* and *Glyma.13G195100*, encoding Toll Interleukin Receptor–nucleotide-binding–leucine-rich repeat resistance proteins were identified as candidate resistance genes by quantitative real-time polymerase chain reaction and sequence analysis. The two closely linked genes work together to cause the phenotypic segregation as a single Mendelian gene. These results will facilitate marker-assisted selection, gene cloning and breeding for the resistance to SC20.

Communicated by Henry T. Nguyen.

✉ Junyi Gai
sri@njau.edu.cn

¹ Soybean Research Institute, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

² National Center for Soybean Improvement, Ministry of Agriculture, Nanjing 210095, Jiangsu, China

³ Key Laboratory of Biology and Genetic Improvement of Soybean, Ministry of Agriculture, Nanjing 210095, Jiangsu, China

⁴ National Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

⁵ Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

Introduction

Soybean [*Glycine max* (L.) Merr] is an important crop cultivated in more than 100 countries as well as a major source of protein and oil for human consumption and animal feed. However, *Soybean mosaic virus* (SMV) is formidable threat to the flourishing soybean production in the world, particularly in China (Cheng et al. 1982) since SMV is an existing problem in three major soybean cultivating areas, i.e., Northeast China, Huang-Huai Valleys and Southern China. Outbreak of SMV greatly decreases the yield and quality of soybean, with yield losses ranging from 50 to 80% (Arif and Hasan 2002).

To cope with the threatening of SMV to Chinese soybean production, a research program was initiated at the National Center for Soybean Improvement (NCSI) on establishment of the nation-wide SMV strain identification system for

further study on the pathogen and breeding soybeans resistant to the pathogen nation-wide. A total of country-wide 22 SMV strains, designated as SC1–SC22 were identified based on their responses to ten soybean differentials along with their distribution and domination in different areas in China (Wang et al. 2003, 2014a, 2005; Guo et al. 2005; Zhan et al. 2006; Li et al. 2010). Among the strains, SC20 widely distributed specifically in Southern China, including five provinces namely Shanghai, Hunan, Fujian, Guangxi and Guangdong where both vegetable soybean and dry-seed soybean are produced.

The most economical and eco-friendly method to control the SMV is to identify resistance genes in *Glycine* species to be used for breeding-resistant varieties. Qihuang-1 was identified as an excellent resistance source to SC20 as well as other strains such as SC7, SC11, SC14, etc (Li et al. 2006, 2010; Bai et al. 2009; Ma et al. 2011), but the genetic and molecular basis of the resistance is still to be revealed.

Breeding for resistance to SMV has been difficult requiring a large amount of artificial inoculation test for the resistance in greenhouses or fields, which is time-consuming and labour-intensive. Molecular markers are powerful tools in tagging the important QTLs/genes in plants (Mohan et al. 1997). Once the tightly linked markers are identified, they can be used to develop marker-assisted selection (MAS) strategy in breeding for SMV resistance.

Recently, soybean breeding programmes using MAS has become sophisticated with the availability of reasonably dense molecular marker linkage map (Song et al. 2004, 2010), and the association of markers to known genes has been pursued by many researchers. Previously, identified and localized SMV-resistance loci from same/different backgrounds are valuable information that soybean breeders can utilize to exploit SMV resistance in a more efficient manner. To date, major SMV-resistance loci have been genetically mapped to Chromosomes 2 (LG D1b), 6 (LG C2), 13 (LG F) and 14 (LG B2). These chromosomes also carry the majority of the resistant gene analogue (RGA)-like genes present in soybean (Hayes et al. 2004; Jeong and Maroof 2004; Moon et al. 2009; Yang and Gai 2010; Suh et al. 2011; Wang et al. 2011a). The molecular markers associated with SMV-resistance genes on these chromosomes were potential for further studying the resistances to diverse strains in a same accession or different backgrounds.

In the United States (US), the symbol *Rsv* was initially used to designate an allelomorph series involved in SMV resistance. The researchers have identified three genes (*Rsv1*, *Rsv3* and *Rsv4*) resistant to G1–G7 SMV strains (Yu et al. 1994; Hayes et al. 2004; Jeong et al. 2002; Jeong and Maroof 2004; Maroof et al. 2010). Among them, *Rsv1* locus represents the most complicated one. One of the difficulties is that *Rsv1* is located in a genomic region that is rich in *R* genes, a mixed cluster of 14 nonTIR–NBS–LRR genes (nTNL) and

5 TIR–NBS–LRR (TNL) genes are presented in this region (Hayes et al. 2004; Gore et al. 2002). These *R* genes confer resistances to various pathogens. Such as the *Rsc* genes confer resistances to the strains of NCSI SMV strain system from China (Li et al. 2006; Fu et al. 2006; Yang and Gai 2010; Wang et al. 2011a), *Rpv1* confers to peanut mottle virus (Gore et al. 2002), *Rps3* confers resistance to *Phytophthora sojae* (Diers et al. 1992), *Rpg1* confers resistance to bacterial blight (Ashfield et al. 1998, 2004), *Rcs* confers resistance to *Cercospora* leaf spot (Pham et al. 2015).

In China, a number of SMV-resistance genes from different resistant sources have been reported. Among those, the genes *Rsa*, *Rn1*, *Rn3*, *Rsc7*, *Rsc8*, *Rsc9*, *Rsc13* and SC18-resistance gene from the cultivar Kefeng-1 were mapped on chromosome 2 (Wang et al. 2004, 2011b; Fu et al. 2006; Guo et al. 2007; Yan et al. 2015; Li et al. 2015). The gene *Rsc15* from RN-9 was mapped on chromosome 6 (Yang and Gai 2010). The gene *Rsc14Q*, *Rsc11*, *Rsc12*, *Rsc3Q* and another SC18-resistance gene from Qihuang-1 and Qihuang-22, *Rsc-pm* and *Rsc-ps* from PI96983 were mapped on chromosome 13 (Li et al. 2006, 2015; Ma et al. 2010, 2011; Yang et al. 2013; Zheng et al. 2014). And the gene *Rsc4* from Dabaima was mapped on chromosome 14 (Wang et al. 2011a). The above-mentioned SMV-resistance genes (except for *Rsc15*) in China were mapped in close proximity to *Rsv* genes reported in the US, but the relationship between *Rsv* and *Rsc* genes remains unknown.

The genome-wide scanning approach has been extensively used to map the SMV-resistance genes. But if the chromosomes harboring the genes which are previously known, a partial genome mapping strategy (or limited mapping strategy) can be used for delimiting the genome regions through choosing markers on the target chromosomes or genome segments using bulked segregant analysis (BSA, Michlemore et al. 1991).

In addition, genomic applications in soybean have become more standard with the availability of whole genome sequence (WGS) (Schmutz et al. 2010). The WGS provided the basis for the development of new markers and identification of candidate genes. Recent technological development in Next-generation sequencing (NGS) provides ability to re-sequence cultivars of interest at relatively low cost. Apart from the functional marker development, the sequence data are useful to know the structural and allelic variation between the genes (Tollenaere et al. 2012).

In the present study, we aimed at revealing the inheritance of resistance to the SMV strain SC20 in Qihuang-1 using Mendelian genetic populations, locating the region of the resistance gene(s) using the partial genome mapping strategy, narrowing down the region into a small chromosome segment using the derived residual heterozygote population, and finally identifying the putative candidates of the resistance gene(s) based on qRT-PCR and sequence analysis.

Materials and methods

Plant genetic materials

Soybean cultivar Qihuang-1 is resistant (R) to the SMV strain SC20 and NN1138-2 is susceptible (S) to this strain. The F₁, F₂ and F_{2:7}-derived RIL (recombinant inbred line) populations obtained from the cross of Qihuang-1 × NN1138-2 along with the parents were used to explore the inheritance of resistance. To map the resistance gene, the RIL population and segregating populations were used. In preparing the RIL population, the F₁ plants were self-pollinated to produce an F₂ population, from which single seed descent was conducted from the F₃–F₇ and then 178 F_{2:7}-derived lines were established as the RIL population of Qihuang-1 × NN1138-2 at the National Center for Soybean Improvement (NCSI), Nanjing agricultural university (NAU), China.

The residual heterozygote-derived population with 346 plants was established from the selfed seeds of four residual heterozygous (RH) individuals (RH51-1, RH51-2, RH51-3 and RH51-4). These four RH individuals were identified from the line RIL51 of the RIL population and carried a heterozygous segment extending from Satt114 to Sct_033 that includes the target gene region (Table 1). The two parental lines and mapping populations were planted in plastic pots (ϕ 20 × h 20 cm) and grown in aphid-free greenhouses at 25 °C at the NAU Pailou experimental station, Nanjing, China.

The SMV isolate, phenotypic assessments and genetic analysis

A single SMV isolate of the strain SC20, provided by the NCSI, was used in this study. Plants were mechanically inoculated with SMV when the unifoliate leaves began to unfold. The inoculum of the virus was obtained in the leaves of the susceptible cultivar, NN1138-2. It was prepared by grinding infected fresh leaves of NN1138-2

in 0.01 mol/l sodium phosphate buffer (approximately 3–5 ml/g leaf tissue, pH 7.2) using a mortar and pestle. A small amount of 600-mesh carborundum powder was added to the inoculum as an abrasive. Infection responses were observed weekly during the 35 days after inoculation and when the symptoms (mosaic, necrosis, and symptomless) appeared stable on leaves emerging above the inoculated leaves by visual assessment. The number of infected plants vs. total plants per pot were all recorded, and the incidence rates (ratio of infected to total plants) of each inoculated line were calculated. The resistance and susceptible plants were classified based on the following criteria. Plants failed to develop visible symptoms and were similar in appearance to uninoculated plants or plants with mild necrotic streaks or necrotic local lesions only on inoculated leaves or plants with very mild mosaic symptom but incidence rate less than 10% were designated as resistance. Plants developed mosaic symptoms, with or without associated necrosis or plants with systemic necrosis were considered as susceptible.

The segregation patterns of phenotypes in the mapping population were tested for the goodness of fit to Mendelian segregation ratio using Chi-square criterion.

DNA extractions, SSR marker choices and preliminary mapping strategy

The genomic DNA was extracted from the young leaves by the cetyltrimethyl ammonium bromide (CTAB) method (Doyle 1990) and was stored at – 40 °C. We used the partial genome mapping strategy to locate the SC20-resistance gene. The process was involved in four steps.

First step, 26 SSR markers reportedly linked to several SMV-resistance loci were selected from the previous studies (Hayes et al. 2004; Fu et al. 2006; Jeong and Maroof 2004; Moon et al. 2009; Yang and Gai 2010; Suh et al. 2011; Wang et al. 2011a). All the markers used were tested for polymorphism between the parents, from which eight polymorphic

Table 1 The marker genotypes of the residual heterozygote individuals

Serial no.	Individual	Marker					
			Satt234	Satt114	SSR_13_1099	SSR_13_1185	Sct_033
1	RH51-1	B	H	H	H	H	A
2	RH51-2	B	H	H	H	H	A
3	RH51-3	B	H	H	H	H	A
4	RH51-4	B	H	H	H	H	A

Marker the markers are linked with SC20-resistance gene and arranged according to their linkage order, *SSR_13_1099* BARCSOYSSR_13_1099, *SSR_13_1185* BARCSOYSSR_13_1185, *A* the homozygous Qihuang-1 chromosome segment, *B* the homozygous NN1138-2 chromosome segment, *H* the heterozygous Qihuang-1/NN1138-2 chromosome segment

SSR markers (Satt266, Satt286, Satt079, Sat_263, Satt114, Satt 297, Sct_033 and Satt687) were identified.

Second step, using the BSA method, Satt 266, Sat_263, Satt114 and Satt687 out from the eight polymorphic SSR markers were employed to identify any marker linked to the SC20-resistance gene. In BSA analysis, based on the phenotyping data of 178 RILs, 16 RILs resistant and 16 RILs susceptible to SC20 were selected for the construction of resistant bulk (R-bulk) and susceptible bulk (S-bulk) by mixing equimolar concentration of DNA samples from the resistant and susceptible RILs, respectively. The polymerase chain reaction (PCR), polyacrylamide gel electrophoresis (PAGE) and silver staining were following the method used by Ma et al. (2011).

Third step, the marker Satt114 found to link to the resistance gene was applied to the full mapping population for confirming the association between the marker and the resistance gene.

Fourth step, additional seven known polymorphic SSR markers in 10 Mb windows both upstream and downstream of Satt114 (27718778) from the region of Satt516 (22489622) to Satt490 (36699478) on Chromosome 13 were screened in 178 RILs.

The resulted data were used to construct a linkage group using Join Map 4.0. The distances between markers and the resistance gene were calculated using Join Map 4.0 linkage analysis software and then transformed into Kosambi genetic distances (Kosambi 1943). The markers were assigned to linkage groups (LGs) by using a minimum logarithm of odds (LOD) likelihood score of 4.0 (Van Ooijen 2006).

Targeted marker creation for fine-mapping the resistance gene

The first round of fine-mapping was performed with the SSR markers of BARCSOYSSR_13_1114, BARCSOYSSR_13_1136 and BARCSOYSSR_13_1140, which were previously reported as polymorphic and existed in the target region (Zheng et al. 2014) on the RH population.

The second round, 83 molecular markers, including 22 presence/absence variation (PAV) markers (Wang et al. 2014b), 5 express sequence tag (EST) SSR (Shirasawa et al. 2014), 42 new SSR and 14 Insertion and Deletion (Indel) markers around the target region were used for the RH population. A DNA segment between BARCSOYSSR_13_1140 and BARCSOYSSR_13_1185 on Chromosome 13 of Williams 82 soybean was identified using a ‘BLAST genome’ search in the Williams 82 soybean reference genome *GlymaWm 82.a2.v1* (<http://soybase.org>). The 42 SSR primers were designed using BatchPrimer3 v1.0 software (<http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi>) (You et al. 2008). The SSR markers containing ideal repeat units of 2–6 nucleotides only were selected. The standard based on the fixation of minimum SSR length was defined as five reiterations for each repeat unit. The complex SSR types and mononucleotide repeats were omitted. A guideline for designing SSR primers was explained by Dutta et al. (2011). Although 14 InDel markers were developed by comparing the sequences of PCR products amplified from Qihuang-1 and NN1138-2 using primers designed from the soybean genome sequence in the target interval. The primers were designed using the Primer Premier 5.0 software (Premier Biosoft International, Palo Alto, CA). All the 83 markers were screened for polymorphism between the parents, the polymorphic markers were tested in the entire RH population and the linkage analysis was done.

Candidate gene prediction and quantitative real-time PCR (qRT-PCR)

Candidate genes in the target region were predicted using the SoyBase, Williams 82 soybean reference genome *GlymaWm 82.a2.v1* (<http://soybase.org>) (verified 14th March, 2016). The qRT-PCR was conducted to obtain expression profiles of candidate genes. Gene-specific primers for qRT-PCR were designed from the respective gene sequences using Primer Premier 5.0 software (Table 2). Tubulin was used as an internal reference gene control. Primary leaves

Table 2 A list of primer sequences used in expression profiling and sequence analysis

Serial no.	Gene name	Primer sequence (5′–3′)	Primer sequence (5′–3′)	Annealing temperature	Size (bp)
1	Glyma.13G194700 ^a	GATTGAGAAGATTGTGGAGGAC	CAGCAAGGGATGTGGCA	58	128
2	Glyma.13G195100 ^a	GGAAAATATGGAACTGTGCG	TTACGTCAAATCCCCTTCC	58	113
3	Glyma.13G195200 ^a	TCCTGCATAGAGCAGTTGCC	ACAAGTGCAGCGGTGTCATA	58	117
4	Tubulin ^a	TGACCAAGACCCCAAAGG	GATGTGGAAAGCCAAAACC	58	148
5	Glyma.13G194700 ^b	GAATTAGCGAAGAAGTAATCG	CTCAGCTTCATTACACTCCAC	56	3600
6	Glyma.13G195100 ^b	GAATTAGCGAAGAAGTAATCG	GGAGGGGAGGAAAGTAGGCT	56	3900

^aPrimer sequences used for expression profiling analysis

^bPrimer sequences used for sequence analysis

of the two parental lines were inoculated with SC20 and buffer (as a mock inoculation), respectively, and kept in a growth chamber at 25 °C. The total RNA was isolated from the leaf at time points at 0, 1, 2, 4, 8, 12, 24, 48, and 72 h post inoculation. Equal amounts of total RNA (25 µg per reaction) were treated with DNase I (Promega) to remove contaminating DNA, and the resulting DNA-free RNA (5 g per reaction) was used for the first strand cDNA synthesis (TaKaRa). The cDNA was diluted to tenfold before being used as the template for qRT-PCR. The qRT-PCR mixture in a 20-µl final volume contained 10 µl of 2× SYBR Green PCR Master Mix (TaKaRa), 0.8 µl of each primer, 2 µl of template (10× diluted cDNA from samples) and 6.4 µl of sterile distilled water. The thermal conditions are as the following: 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s and a final step at 60 °C for 1 min. All reactions were performed in three times in 96-well reaction plates using Roche Light-Cycler[®] 480 (Roche Diagnostics, Indiana, USA). Three independent replications were performed. The expression was quantified using the relative quantification ($2^{-\Delta\Delta C_t}$) method and the data were compared with the respective internal controls.

Amplification and sequencing of resistance gene candidates

The gene-specific primers were designed based on the 5'-untranslated regions (UTR) and 3'-UTR of the gene models *Glyma.13G194700* and *Glyma.13G195100* in Williams 82 soybean reference genome *GlymaWm 82.a2.v1*. The primer pairs were designed using Primer Premier 5.0 software (Premier Biosoft International, Palo Alto, CA) and then synthesized (Table 2). The full-length coding sequences (CDS) of each gene were amplified using cDNA from Qihuang-1 and NN 1138-2. PCR products were purified and ligated into the pMD19-T vector (Takara) overnight at 4 °C according to the manufacturer's protocol. The ligated product was transformed into DH5α competent cells and plated into a LB agar medium containing 100 µg/ml ampicillin, 50 mg/ml X-gal and 50 mg/ml IPTG and grown overnight

at 37 °C. The positive clones (white colonies) were screened through PCR and then sequenced by Invitrogen Biotechnology Co., Ltd (Shanghai, China). The amino acid prediction was done using online program softberry FGENESH (<http://www.softberry.com>) and the sequence was used for confirming the motifs in comparison to the NCBI reference protein database using BLASTP (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). For protein annotation, InterPro (<http://www.ebi.ac.uk/interpro/>) was used. Nucleotide and protein alignments were performed using Clustal omega (<http://www.ebi.ac.uk/Tools/msa/clustalo>) with default parameters.

Results

Inheritance of resistance to SC20

The resistant parent Qihuang-1 plants did not develop any disease symptoms to SMV strain SC20 even after 35 days post inoculation (dpi), while the susceptible parent NN1138-2 plants started developing mosaic symptoms from 10 to 35 dpi. All F₁ plants derived from the cross between Qihuang-1(R) and NN1138-2 (S) were resistant to the SMV strain SC20, indicating that SC20 resistance is dominantly inherited. The observed segregation for SMV resistance in the F₂ generation fitted the expected phenotypic segregation ratio 3R:1S. Furthermore, disease response in the RIL population fitted the expected genotypic (also phenotypic) ratio 1R:1S (Table 3). These results advocate that a single dominant gene is involved in the resistance to the SMV strain SC20 in Qihuang-1. This result was further validated by the phenotypic ratio 3R:1S in the residual heterozygote-derived population (Table 3).

Preliminary mapping of the SC20-resistance gene

A total of 26 SSR markers represented several SMV-resistance loci were deployed to identify the marker(s) linked with the SC20-resistance gene. The BSA analysis showed that the SSR marker Satt114 exhibited polymorphisms

Table 3 Genetic of resistance to Soybean mosaic virus (strain SC20) in Qihuang-1

Serial no.	Parent or progeny	No. plants (lines)		Total	Expected ratios	χ^2	<i>p</i> value
		R	S				
1	Qihuang-1	20	–	20			
2	NN1138-2	–	20	20			
3	F ₁	10	–	10			
4	F ₂	73	47	120	3:1	1.83	0.17
5	RILs	95	83	178	1:1	0.36	0.54
6	The residual heterozygote-derived population	266	80	346	3:1	0.65	0.42

between the resistant and susceptible parents, as well as between the resistant and susceptible bulk groups. To confirm the reliability of linkage between the marker Satt114 and the resistance gene, we assayed the 178 lines of the Qihuang-1 × NN1138-2 RIL population using the marker Satt114. The linkage analysis results revealed that Satt114 really linked to the resistance gene from Qihuang-1, exhibiting genetic distances of 7.0 cM.

To further construct the linkage group and map the resistance gene, additional seven polymorphic SSR markers in the target region were added for further linkage analysis. The resulted data were used to construct a linkage group using Join Map 4.0. The total length of the linkage group covered by eight markers was 62.0 cM and the average genetic distance was 7.8 cM per marker. The resistance gene was positioned to an interval on Chromosome 13 between BARCSOYSSR_13_1099 and BARCSOYSSR_13_1185. The genetic distances for the resistance gene to the two respective markers were 3.9 and 3.8 cM, and those to all the eight markers are presented in Table 4.

Fine-mapping for delimitation of the candidate genomic region

To narrow the region harboring the SC20-resistance gene, a large RH population with 346 plants was adopted to further map the targeted gene. At first, to verify our preliminary mapping results and confirm the map location of the SC20-resistance gene, three SSR markers BARCSOYSSR_13_1099, BARCSOYSSR_13_1185 and Sct_033 were screened out. The linkage analysis using the three markers confirmed that the resistance gene was located between BARCSOYSSR_13_1099 and BARCSOYSSR_13_1185. Thus, three additional polymorphic SSR markers, BARCSOYSSR_13_1114, BARCSOYSSR_13_1136 and

BARCSOYSSR_13_1140 between the region of BARCSOYSSR_13_1099 and BARCSOYSSR_13_1185 were selected and evaluated.

The linkage analysis using the three new markers further revealed that the resistance gene was located between two flanking markers, namely BARCSOYSSR_13_1140 and BARCSOYSSR_13_1185 (Fig. 1a) and 22 recombinants were identified to have the resistance gene between them.

To identify additional markers in the delimited target region that are polymorphic between Qihuang-1 and NN1138-2, the 22 PAV and 5 EST markers located in this region (Wang et al. 2014b; Shirasawa et al. 2014) as well as the 56 new markers (42 SSR and 14 Indel markers) based on the sequence data of Williams 82 (<http://soybase.org/>) were developed and tested. A total of six markers, gm-ssr_13-9, gm-ssr_13-14, gm-indel_13-3, gm-ssr_13-28, gm-indel_13-12 and GMES 5258 (Table 5), were identified as polymorphic in this screening. Consequently, the recombinants of the RH population were then genotyped with these markers. Eventually, SC20-resistance gene was delimited to a 0.4 cM interval flanked by gm-ssr_13-14 and gm-indel_13-3 using Join Map 4.0. The high-resolution genetic map of SC20-resistance gene is presented in Fig. 1. The phenotypes and genotypes of 22 recombinants showing that the recombinant breaking points was presented in Table 6, from which the SC20-resistance gene located also between gm-ssr_13-14 and gm-indel_13-3. Here combining the phenotypic and genotypic information from the 22 lines in Table 6, the possible resistance gene in recombinants with R phenotype might locate in the region of gm-ssr_13-14, gm-indel_13-3 and gm-ssr_13-28 (representative line 181), while the possible susceptible gene in recombinants with S phenotype might locate in the region of gm-ssr_13-14 and gm-indel_13-3 (representative line 108) and possible around gm-ssr_13-14 (representative line 35). Therefore, for a locus

Table 4 Position of SC20-resistance gene and their relative markers on chromosome 13 (LG F) in Qihuang-1 × NN1138-2 in the preliminary mapping experiment

Serial no.	Marker	Position on Chr.13 (cM)	Neighboring distance (cM)	Physical location (starting position, bp)
1	Satt516	0.0	–	22,489,622
2	Sat_234	17.8	17.8	27,656,895
3	Satt114	24.5	6.7	28,912,864
4	SSR_13_1099	27.6	3.1	29,609,537
5	SC 20 resistance gene	31.5	3.9	–
6	SSR_13_1185	35.3	3.8	31,348,154
7	Sct_033	39.6	4.3	31,951,960
8	Satt313	50.2	10.6	35,508,621
9	Satt490	62.0	11.8	36,699,478

Physical location of the markers is based on SoyBase database, Williams 82 soybean reference genome annotation

Glyma Wm 82.a2.v1 (accessible at <http://soybase.org/>) (verified 14th March 2016)

SSR_13_1099 BARCSOYSSR_13_1099, SSR_13_1185 BARCSOYSSR_13_1185

Fig. 1 Map position of resistance gene in soybean cultivar Qihuang-1 conferring resistance to soybean mosaic virus strain SC20 on chromosome 13 (LG F). **a** Linkage map based on three markers in a segregating population of 346 individuals derived from RH (residual heterozygote) plants, **b** linkage map based on twelve markers in a segregating population of 346 individuals derived from RH plants. Map distances (cM) between neighboring markers are on the left with marker names on the right

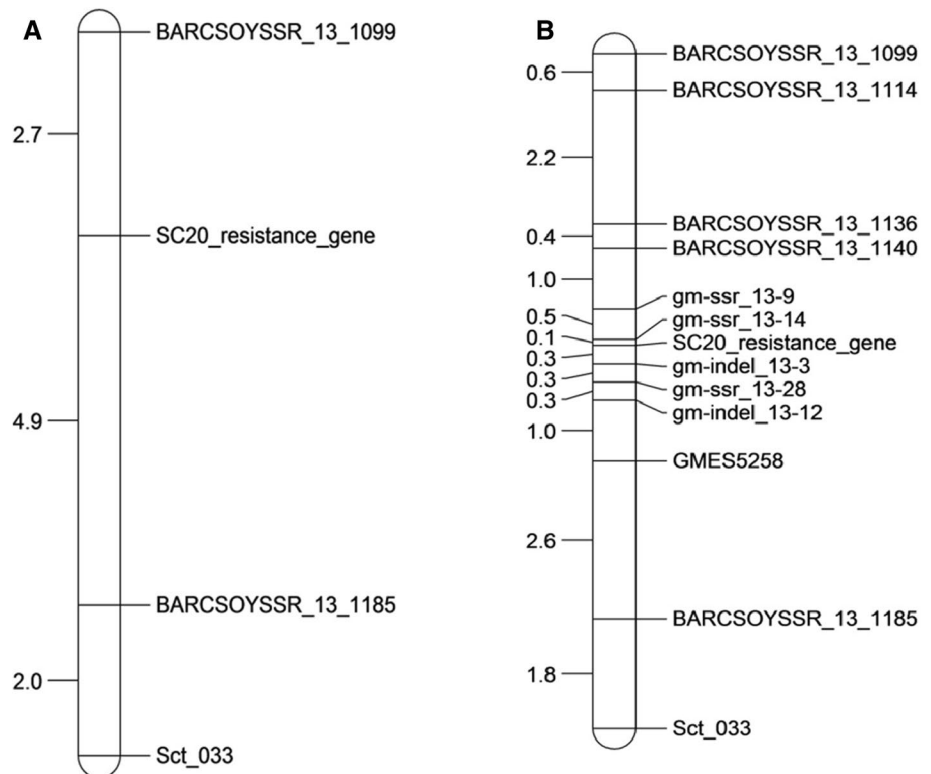


Table 5 The sequence information of molecular markers used for delimiting the candidate genomic region in the second round fine-mapping experiment

Serial no.	Marker	Primer sequence (5'–3')		Genomic position ^d (starting position of the marker, bp)	Annealing temperature	Size (bp)
		Forward	Reverse			
1	gm-ssr_13-9 ^a	GGGCCCGTTTGTAAAGG	GAAAACGGGTTTTGTAGA TTAGT	30,684,564	55	246
2	gm-ssr_13-14 ^a	ACAAAACGACAATTTTACCCA	CGACCGGATACTGTAAGA	30,795,177	55	181
3	gm-indel_13-3 ^c	ACAAATCTTAAGAACAGG GAAC	ACATCCTCCCATCCATATCT	30,875,000	55	246
4	gm-ssr_13-28 ^a	TGATAGATTTCCACCAAC	ATGGACACCACTCTAAAG AATC	30,899,513	55	225
5	gm-indel_13-12 ^c	GGTGAGTTTGGTCGATAATGT	AGACTCGCGTAAGGTGTTCT	31,003,191	56	484
6	GMES5258 ^b	CGACAGCCAAATCCAAAAT	GGCATCTGCAACAACCTCTGA	31,181,520	55	199

^aSSR marker

^bEST–SSR marker

^cIndel marker

^dGenomic position of each marker was determined by BLASTN based on SoyBase Williams 82 soybean reference genome Glyma Wm 82.a2.v1 (<http://soybase.org>)

with both resistance and susceptible alleles, it must be in the overlapped region, i.e. in or between gm-ssr_13-14 and gm-indel_13-3.

Identification of candidate genes in the delimited region

To locate the resistance gene in the soybean genome sequence, BLASTN searches for the sequences of the two markers tightly linked to the SC20-resistance gene were

Table 6 Phenotypes and genotypes of recombinant plants showing the recombinant breaking points

S.NO.	Phenotype	BARCSOYSSR		gm-ssr	gm-ssr	gm-ssr	gm-indel	gm-ssr	gm-indel	GMES 5258	BARCSOYSSR	
		13	1140	13-9	13-14	13-3	13-28	13-12	13	1185		
Qihuang-1	R	a	a	a	a	a	a	a	a	a	a	a
NN 1138-2	S	b	b	b	b	b	b	b	b	b	b	b
77	R	a	a	a	a	a	a	a	a	a	a	h
73	R	a	a	a	a	a	a	h	h	h	h	h
267	R	a	a	a	a	a	a	h	h	h	h	h
23	R	a	a	a	h	h	h	h	h	h	h	h
21	R	a	h	h	h	h	h	h	h	h	h	h
80	R	a	h	h	h	h	h	h	h	h	h	h
178	R	a	h	h	h	h	h	h	h	h	h	h
143	R	h	h	h	h	h	h	h	h	a	a	a
328	R	h	h	h	h	h	h	h	h	a	a	a
55	R	h	h	h	h	h	h	a	a	a	a	a
181	R	h	h	h	h	h	h	b	b	b	b	b
214	R	h	h	h	h	h	a	a	a	a	a	a
56	S	h	h	h	b	b	b	b	b	b	b	b
312	S	h	h	h	b	b	b	b	b	b	b	b
93	R	h	h	a	a	a	a	a	a	a	a	a
99	R	h	a	a	a	a	a	a	a	a	a	a
280	S	h	b	b	b	b	b	b	b	b	b	b
108	S	b	b	b	b	h	h	h	h	h	h	h
35	S	b	b	b	h	h	h	h	h	h	h	h
69	R	b	b	h	h	h	h	h	h	h	h	h
173	R	b	b	h	h	h	h	h	h	h	h	h
166	R	b	h	h	h	h	h	h	h	h	h	h

S. no. serial number of the recombinant plants, *R* resistant, *S* susceptible, *a* alleles from the resistant parent, *b* alleles indicates from the susceptible parent, *h* alleles from heterozygote

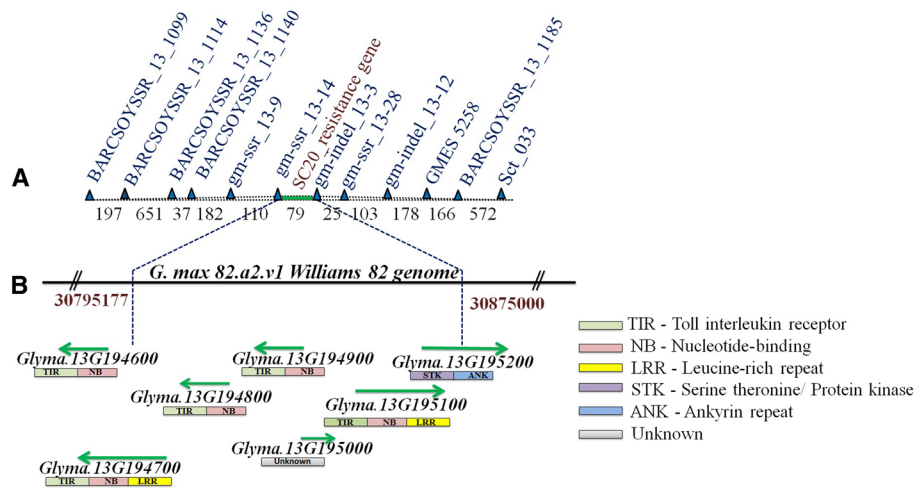


Fig. 2 A physical map of the SC20-resistance gene between the markers gm-ssr_13-14 and gm-indel_13-3 and the candidate region of the SC20-resistance gene. **a** The map position of the SC20-resistance gene on Chr. 13 (the numbers below the map are the estimated relative physical distances in kb based on the reference sequence). **b**

The candidate region of the SC20-resistance gene (a total of seven putative genes were existed in the target genomic region, according to the Williams 82 soybean reference genome annotation *Glyma* Wm 82.a2.v1 (accessible at <http://soybase.org/>) (verified 14th March 2016)

performed initially against the soybean genome sequence of Williams 82.a2.v1 (<http://soybase.org/>). The physical distance of the region between markers gm-ssr_13-14 and gm-indel_13-3 at nucleotide positions 30,795,177 and 30,875,000, respectively, is approximately 79 kb (Fig. 2).

Seven putative genes were existed in the corresponding genomic region. The domain of *Glyma 13G195000* was unknown, and *Glyma 13G195200* belongs to the domain of serine threonine/protein kinase/ankyrin repeat. The rest five genes known to be a family of NBS–LRR genes are the most common disease-resistance (*R*) genes in plants. Among these, two gene sequences (*Glyma.13G194700* and *Glyma.13G195100*) encoding full-length of the Toll interleukin receptor (TIR)–nucleotide-binding (NB)–leucine-rich repeat (LRR) proteins are highly homologous (99% amino acid identity). Another three gene sequences (*Glyma.13G194600*, *Glyma.13G194800* and *Glyma.13G194900*) were significantly shorter, not encoding a LRR protein, sharing 54–58% amino acid identity, and referred as truncated gene products. TIR and NB domains were not sufficient to confer disease resistance; therefore, truncated genes *Glyma.13G194600*, *Glyma.13G194800*, and *Glyma.13G194900* were eliminated for further study. In summary, we focused only on three candidates (*Glyma.13G194700*, *Glyma 13G195100* and *Glyma 13G195200*) to identify a SC20-resistance gene. The functions of these genes were tested further using qRT-PCR analysis.

Expression profiling for identifying resistance genes from the three candidates

The expression profiling of the three candidates was investigated for soybean leaves inoculated with strain SC20, individually, using qRT-PCR analysis. Out of which, two genes, i.e., *Glyma.13G194700* and *Glyma.13G195100* had functional implication after SC20 infection and differentially expressed in Qihuang-1 and NN1138-2. However, the gene *Glyma.13G195200* encoding serine threonine/protein kinase/ankyrin repeat with relatively low expression level was detected. Following inoculation, *Glyma.13G194700*

transcript remained stable expression before 8 hours post inoculation (hpi), then reached maximum expression at 8 hpi by approximate 3.70-fold. From 8 to 72 hpi, the transcript levels decreased steadily and was followed by a slight increase at 48 hpi (Fig. 3). A similar expression pattern was also observed in gene *Glyma.13G195100* (Fig. 3). However, only few changes were seen in *Glyma.13G195100* expression pattern comparing to *Glyma.13G194700* (Fig. 3). These results suggested that the expression level of *Glyma.13G194700* and *Glyma.13G195100* among the three candidate genes was changed by virus infection and could be involved in disease-defense mechanisms.

Sequence analysis for allelic difference of the two resistance genes

To differentiate the resistant and susceptible alleles, the two identified candidate genes were completely sequenced and investigated for the presence of the characteristic motifs and amino acid changes. Using gene-specific primer pairs, the two candidate gene sequences were amplified from resistant to susceptible parents. Within the coding region of the first gene *Glyma.13G194700*, Qihuang-1 showed 98% identity to NN1138-2, and the identified SNPs and Indels that causes changes in amino acid sequences between the two parents showed 97% amino acid identity (Fig. 4). Thus, the coding sequences were different between Qihuang-1 and NN1138-2, including the difference in insertions of 126-nt and 3-nt of the Qihuang-1 gene relative to NN1138-2 gene, (positions – 1056 and – 3249, respectively) and 57-nt deletion in Qihuang-1 (position-3287), as well as 63 SNPs identified between them in the coding region. Furthermore, the amino acid changes were observed at TIR domain (2 changes), NB domain (42 changes), and LRR domain (3 changes) (Fig. 5).

Next, for the gene *Glyma.13G195100*, the comparison between the two parents revealed that both parents had 99% identity at nucleotide and amino acid level. The SNPs and Indels identified in NN1138-2 comparing to the sequence of Qihuang-1 (Fig. 4), resulted in amino acid changes between the two parents (Fig. 5). Assessment of coding sequence

Fig. 3 Expression profiling of candidate genes *Glyma.13G194700* and *Glyma.13G195100*. Y-axes indicate the ratios of relative fold expression levels between samples infected with soybean mosaic virus

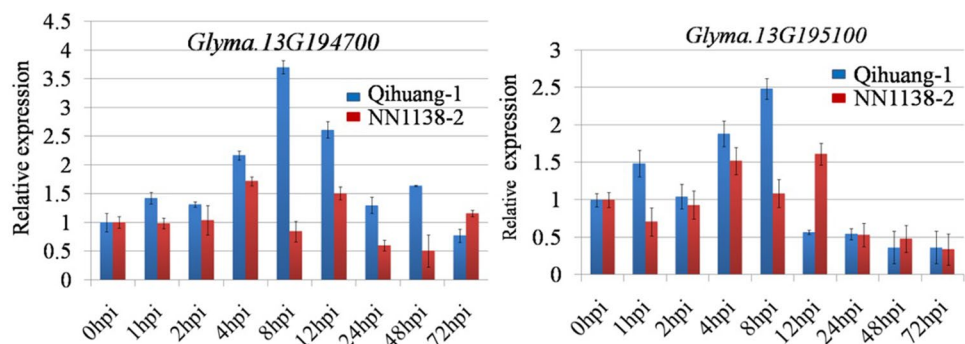
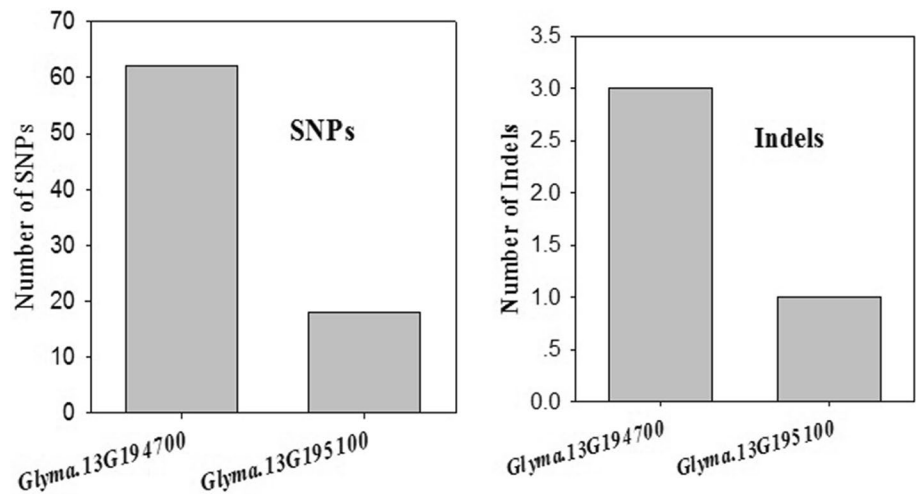


Fig. 4 Number of SNPs and Indels between the two candidate genes in parental lines of Qihuang-1 and NN1138-2



revealed difference between Qihuang-1 and NN1138-2. This difference included insertions of 43-nt in the Qihuang-1 gene relative to NN1138-2 gene, (position-1158), and 17 SNPs identified between the two parents' coding region. Additional amino acid changes were detected at NB domain where 25 changes were identified between the two parents (Fig. 5).

In summary, both genes had amino acid differences in the two parents. These amino acid changes were identified in different parts of TIR–NB–LRR domain (Fig. 5). It supports the notion that the two parents have different alleles and the allelic variations between Qihuang-1 and NN1138-2 may account for the resistance to SC20 in Qihuang-1 or for susceptibility to SC20 in NN1138-2. According to the above results, the two genes, *Glyma.13G194700* and *Glyma.13G195100* are jointly responsible for SC20 resistance. In addition, the two genes might work together closely to cause the phenotypic segregation in a single dominant gene style since they are closely located.

Discussion

The broad resistance spectrum of Qihuang-1 and the effectiveness of RH-derived population in fine-mapping

Qihuang-1 is one of the most important parents with very high frequency in the pedigree of soybean cultivars in China (Xu et al. 2004). Qihuang-1 is resistant to 18 of 22 strains, including SC1–SC8, SC11–SC14, SC16–SC18, and SC20–SC22 (Wang et al. 2014a; Li et al. 2010), and it has been widely used in mapping SMV-resistance genes as well as breeding for SMV-resistant soybean cultivars in China. To date, 92 soybean varieties have been developed from Qihuang-1, among them Qihuang 22 is one of the soybean varieties resistant to SMV. Qihuang-1 along with

its derived varieties has a broad spectrum of SMV resistance, six resistance genes have been mapped or identified, all focused on Chromosome 13 (Fig. 6), while those of the other resistance genes are left for future studies. Qihuang-1 is an important material in studying the SMV-resistance mechanism of many genes clustering on a same chromosome. In addition, there are four strains, SC9, SC10, SC15 and SC19, their resistance genes are not included in Qihuang-1 and are to be transferred to Qihuang-1 for creating a perfect resistance material through using the other resistance sources, such as Kefeng-1, Dabaima and RN-9.

The present study indicated that the resistance to SC20 in Qihuang-1 was due to a dominant Mendelian gene. It was mapped initially to a 1.75 Mb genomic region flanked by SSR markers BARCSOYSSR_13_1099 and BARCSOYSSR_13_1185 on Chromosome 13 using Qihuang-1 × NN1138-2 RIL population. For fine-mapping the resistance gene, an advanced backcrossing materials that are heterozygous at target region and homozygous in the other regions is required. In literature, various advanced populations such as near-isogenic lines (NILs) and chromosome segment substitution lines (CSSLs) can serve fine-mapping Mendelian genes through reducing or eliminating the genetic background noise. However, segregating population or advanced F₂-equivalent population (RH-derived population) or a derived population from a recombinant inbred line has been used more often recently to validate and fine map the QTL/genes (Yamanaka et al. 2005; Tang et al. 2013; Song et al. 2016). For instance, various RH-derived populations have been used effectively in fine-mapping genes of SMV resistance (Ma et al. 2011; Wang et al. 2011b), alkaline salt tolerance (Tuyen et al. 2010), flowering time (Su et al. 2010), sudden death syndrome (Yamanaka et al. 2006) and seed hardness (Hirata et al. 2014) in soybean and grain shape (Liu et al. 2016), chlorophyll content (Jiang et al. 2012) and leaf width (Chen et al. 2012) in rice. In the

Fig. 5 The candidate genes *Glyma.13G194700* (a) and *Glyma.13G195100* (b) deduced amino acid structure of the predicted proteins and alignment of amino acid sequences from Qihuang-1 (resistant) and NN-1138-2 (susceptible). The variation at amino acid level was identified in domains of the TIR–NB–LRR resistance genes protein sequence (note: *Glyma.13G194700* in TIR, NB, LRR domains position between 14–183, 225–445 and 681–922, *Glyma.13G195100* in TIR, NB, LRR domains position between 14–183, 225–445 and 641–922. Domains TIR—green box, NB—red box and LRR—yellow box)

A *Glyma.13G194700*

	<u>TIR</u>	
Qihuang-1	MSKAVSESIDIRAYDVFLSFRGEDTRRSFTGNLYNLEKRGHITFIGDYDFESGEEIKAS	60
NN1138-2	MSKAVSESIGIRAYDVFLSFRGEDTRRSFTGNLYNLEKRGHITFIGDYDFESGEEIKAS	60
Qihuang-1	LSEAIEHSRVFVIVFSENAYSSWCLDELVRI LDPFTEDNHRPVI PVFYVDPDSDVRHQKG	120
NN1138-2	LSEAIEHSRVFVIVFSENAYSSWCLDELVRI LDPFTEDNHRPVI PVFYVDPDSDVRHQKG	120
Qihuang-1	TYGEALAMHERRLNHESYKVMKWRNALRQAATLSGWPFKRGDGYEYELIEKIVEDI SNKI	180
NN1138-2	TYGEALAMHERRLNHESYKVMKWRNALRQAATLSGWPFKRGDGYEYELIEKIVEDI SNKI	180
Qihuang-1	KTSRPVVDRPVGLEYRMLVNLDDATSLAGVHVIGICGMMGIGKTTLARAVVHSAAGHF	240
NN1138-2	KTSRPVVDRPVGLEYRMLVNLDDATSLAGVHVIGICGMMGIGKTTLARAVVHSAAGHF	240
Qihuang-1	DTSCFLGNVRENAMKHGLVHLQQTLLAEIFRENNIRLTSVEQGSLIKKRLHRKRLLLVL	300
NN1138-2	DTSCFLANVRENAMKHGLVHLQQTLLAEIFRENNIRLTSVEQGSLIKKRLHRKRLLLVL	300
Qihuang-1	DDVRELDLRLALVGS PDWFGPGSRVITTRDRQLLKSHGVDRVYEVVLTNGEALLELW	360
NN1138-2	DDVCELDLRLALVGS PDWFGPGSRVITTRDRQLLKSHGVDRVYEVVLTNGEALLELW	352
Qihuang-1	KAFRTVVRVHPDLINKLNRAITFASGIPLALELIGSSLYGRGIEEWESTLDQYEKNPHRDI	420
NN1138-2	-----SSLYGRGIEEWESTLDQYEKNPHRDI	378
Qihuang-1	<u>NB</u>	
NN1138-2	<u>NB</u>	
Qihuang-1	HMAIKISFDALGYLEKEVFLDIACFNGFELAEIEHILGAHHGCC LKIHIGALVEKSLIM	480
NN1138-2	HMAIKISFDALGYLEKEVFLDIACFNGFELAEIEHILGAHHGCC LKIHIGALVEKSLIM	438
Qihuang-1	IDEHGRVQMHDLIQQMGREIVRQESQEHGPKRSRLWATEDIVRVLEDNTGTCKIQSIIID	540
NN1138-2	IDEHGRVQMHDLIQQMGREIVRQESQEHGPKRSRLWATEDIVRVLEDNTGTCKIQSIIID	498
Qihuang-1	FSKSKKVVQWDGMFVKMISLRTLIRKECF SKGPKKLPNSLRVLEWVGYPKSLPSDFK	600
NN1138-2	FSKSKKVVQWDGMFVKMISLRTLIRNECF SKGPKKLPNSLRVLEWVGYPKSLPSDFK	558
Qihuang-1	PEKLAI LKLPYSGEMSL ELPNLFHMRVNLNDRCEFLTRTPDLSGAPI LKELSSVFCENLV	660
NN1138-2	PEKLAI LKLPYSGEMSL ELPNLFHMRVNLNDRCEFLTRTPDLSGAPI LKELSSVFCENLV	618
Qihuang-1	<u>LRR</u>	
NN1138-2	<u>LRR</u>	
Qihuang-1	EIHD SVGFLDKLEIMNFEGCSKLETFPPIKLTSL ESINFSYCSSLVSPFEILGRMENITH	720
NN1138-2	EIHD SVGFLDKLEIMNFEGCSKLETFPPIKLTSL ESINFSYCSSLVSPFEILGRMENITH	678
Qihuang-1	LSLEYTAISKLPNSIRELVRLQSL ELHNCGMLQLPSSIVTLRELEVLSICQCEGLRFSKQ	780
NN1138-2	LSLEYTAISKLPNSIRELVRLQSL ELHNCGMVQLPSSIVTLRELEVLSICQCEGLRFSKQ	738
Qihuang-1	DEDVKNKSLMPSSYLKQVNLWSCSISDELIDTGLAWFANVRKLDLSANNFTILPSCIQE	840
NN1138-2	DEDVKNKSLMPSSYLKQVNLWSCSISDELIDTGLAWFANVRKLDLSANNFTILPSCIQE	798
Qihuang-1	CRLLRKLYLDYCMHLQEIRGIPPNETLSAIRCTS LKDLAVPLESTREGCCRLQLIID	900
NN1138-2	CHLLRKLYLDYCMHLQEIRGIPPNETLSAIRCTS LKDLAVPLESTREGCCRLQLIID	858
Qihuang-1	<u>LRR</u>	
NN1138-2	<u>LRR</u>	
Qihuang-1	DCENLQEIRGIPPSIEFLSATNCRSLTASCRMLLKQELHEAGNKRYSLPGTRIPWFEH	960
NN1138-2	DCENLQEIRGIPPSIEFLSATNCRSLTASCRMLLKQELHEAGNKRYSLPGTRIPWFEH	918
Qihuang-1	CSRGQISIFWFRNKFPVLSLCLAGLMHKHPFGLKPIVSI NGNKMKTEFORRWYFEPFVL	1020
NN1138-2	CSRGQISIFWFRNKFPVLSLCLAGLMHKHPFGLKPIVSI NGNKMKTEFORRWYFEPFVL	978
Qihuang-1	TDHILTFGERQIKFEDNVDEVVSENDWNHVVSVDVDFKWNPTPELVVVRTGLHVIKPKSS	1080
NN1138-2	TDHILTFGERQIKFEDNVDEVVSENDWNHVVSVDVDFKWNPTPELVVVRTGLHVIKPKSS	1037
Qihuang-1	VEDIRFIDPYKPTIL-----	1095
NN1138-2	KQLLALISPFVMNYLVCEDNYYLLSSRKS LWRCK	1071

Fig. 5 (continued)

B *Glyma.13G195100*

	<u>TIR</u>	
Qihuang-1	<u>MSKAVSESIDIRAYDVFLSFRGEDTRRSPTGNLYNLEKRGIIHTFIGDYDFESGEEIKAS</u>	60
NN1138-2	<u>MSKAVSESIDIRAYDVFLSFRGEDTRRSPTGNLYNLEKRGIIHTFIGDYDFESGEEIKAS</u>	60

Qihuang-1	<u>LSEAIERSRVFVIVFSENYASSWCLELVRIIDFTEDNHRFPVIVFYVDFPSDVRHQKG</u>	120
NN1138-2	<u>LSEAIERSRVFVIVFSENYASSWCLELVRIIDFTEDNHRFPVIVFYVDFPSDVRHQKG</u>	120

Qihuang-1	<u>IYGEALAMHERRLNHESYKVMKRNALRQAATLSGWPFKHDGVEYELIEKIVEDI SNKI</u>	180
NN1138-2	<u>IYGEALAMHERRLNHESYKVMKRNALRQAATLSGWPFKHDGVEYELIEKIVEDI SNKI</u>	180

	<u>TIR</u>	<u>NB</u>
Qihuang-1	<u>KISRFVDRPVGLEYRMLVNWLLDATSLAGVHVIGICGMGGIKTFLARAVYHSAAGHF</u>	240
NN1138-2	<u>KISRFVDRPVGLEYRMLVNWLLDATSLAGVHVIGICGMGGIKTFLARAVYHSAAGHF</u>	240

Qihuang-1	<u>DTSCFLGNVRENAMKHGLVHLQQTLLAEIFRENNIRLTSVEQGISLIKRLHRKRLLLFL</u>	300
NN1138-2	<u>DTSCFLANVRENAMKHGLVHLQQTLLAEIFRENNIRLTSVEQGISLIKRLHRKRLLLVL</u>	300

Qihuang-1	<u>DDVRELEDLQALVGSFDWFGPGSRVIITTRDRQLLKAHGVDKVEVEVLNGEALKLLCW</u>	360
NN1138-2	<u>DDVRELEDLQALVGSFDWFGPGSRVIITTRDRQLLKAHGVDKVEVEVLNGEALKLLCW</u>	360

Qihuang-1	<u>KAFRTDRVHPDFISKLNRAIFASGIPLVLELIGSSLYGRGIEEWESTLDQYKPNPHRDI</u>	420
NN1138-2	<u>KAFRTVRVHPDLINKLNRAITFASGI-----IEEWESTLDQYKPNPHRDI</u>	405

	<u>NB</u>	
Qihuang-1	<u>HMALKISFDALSYLEKEVFLDIACFNGFELAEIEHILGAHHGCCLKIHIGALVEKSLIM</u>	480
NN1138-2	<u>HMALKISFDALGYLEKEVFLDIACFNGFELAEIEHILGAHHGCCLKIHIGALVEKSLIM</u>	465

Qihuang-1	<u>IDEHGCVMQMDLMQQMGREIVRQESQEHGPKRSRLWSTEDIVRVLEDNTGTSKIQSIIL</u>	540
NN1138-2	<u>IDEHGCVMQMDLMQQMGREIVRQESQEHGPKRSRLWSTEDIVRVLEDNTGTSKIQSIIL</u>	525

	<u>LRR</u>	
Qihuang-1	<u>PEKLAIKLPYSGFMSLELPNPLHMRVNLNFDRCFKLTRTPDLGAPILKELSSVFCENLV</u>	660
NN1138-2	<u>PEKLAIKLPYSGFMSLELPNPLHMRVNLNFDRCFKLTRTPDLGAPILKELSSVFCENLV</u>	645

Qihuang-1	<u>EIHDSVGFLDKLEIMNFESCSKLETFPPPIKLTSLLESINFSYCSSLVSFPPEILGRMENITH</u>	720
NN1138-2	<u>EIHDSVGFLDKLEIMNFESCSKLETFPPPIKLTSLLESINFSYCSSLVSFPPEILGRMENITH</u>	705

Qihuang-1	<u>LSLEYTAISKLPNSIRELVRLQSLLEHNCGMVQLPSSIVTLRELEVLISICQCEGIRFSKQ</u>	780
NN1138-2	<u>LSLEYTAISKLPNSIRELVRLQSLLEHNCGMVQLPSSIVTLRELEVLISICQCEGIRFSKQ</u>	765

Qihuang-1	<u>DEDVKNKSLMPSYLRQVNLWSCSISDEFIDTGLAWFANVKSIDLANNFTIILPSCIQE</u>	840
NN1138-2	<u>DEDVKNKSLMPSYLRQVNLWSCSISDEFIDTGLAWFANVKSIDLANNFTIILPSCIQE</u>	825

Qihuang-1	<u>CRLLRKLYLDYCTHLHEIRGIPPNLETLSAIRCTS LKDLAVPLESTREGCCLRQLILD</u>	900
NN1138-2	<u>CRLLRKLYLDYCTHLHEIRGIPPNLETLSAIRCTS LKDLAVPLESTREGCCLRQLILD</u>	885

	<u>LRR</u>	
Qihuang-1	<u>DCENLQEIIRGIPPSIEFLSATNCRSLTASCRMRLLKQELHEAGNKRYSLPGTRIPEWFEH</u>	960
NN1138-2	<u>DCENLQEIIRGIPPSIEFLSATNCRSLTASCRMRLLKQELHEAGNKRYSLPGTRIPEWFEH</u>	945

Qihuang-1	<u>CSRGQISIFWFRNKFVVISLCLAGLMHKHFFGLKPIVSIINGNRMKTEFQRWWFYFEPVL</u>	1020
NN1138-2	<u>CSRGQISIFWFRNKFVVISLCLAGLMHKHFFGLKPIVSIINGNRMKTEFQRWWFYFEPVL</u>	1005

Qihuang-1	<u>TDHILIFGERQIKFEDNVDEVVSENDWNHVVSVDVDFKWNPTPELVVRTGLHVIKPKSS</u>	1080
NN1138-2	<u>TDHILIFGERQIKFEDNVDEVVSENDWNHVVSVDVDFKWNPTPELVVRTGLHVIKPKSS</u>	1065

Qihuang-1	<u>VEDIRFIDPYKPTFL</u>	1095
NN1138-2	<u>VEDIRFIDPYKPTFL</u>	1080

present study, based on the initial mapping, we developed a RH-derived population by selfing RHLs ($F_{2:7}$) to eliminate genetic background noise. This population performed

effectively in delimiting the resistance gene into a 79-kb genomic region.

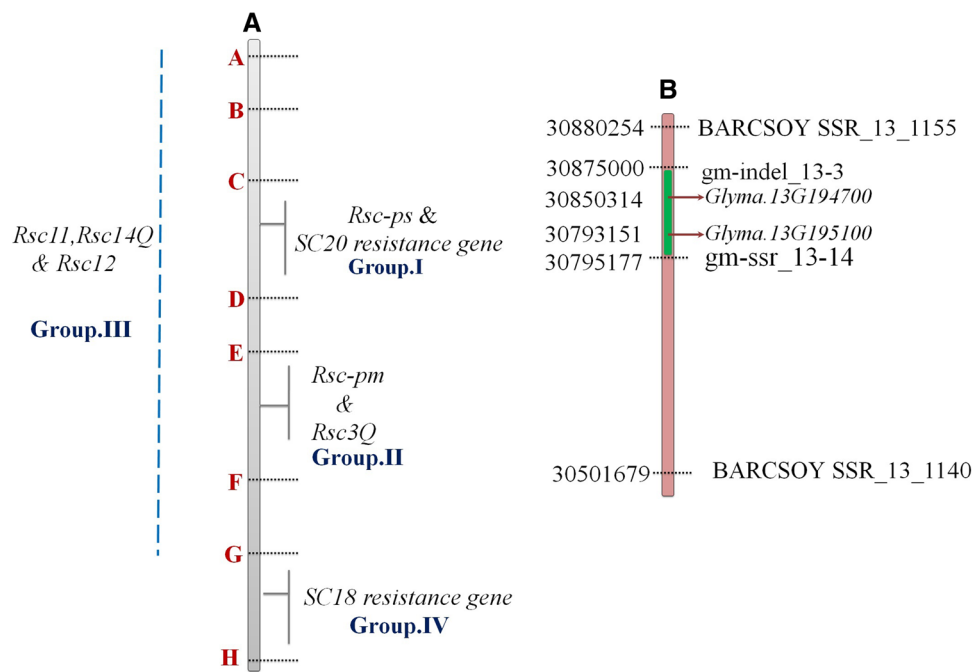


Fig. 6 Integrative views of the SC20-resistance gene location with other previously published SMV resistant genes on chromosome 13. **a** The gene/loci can be classified into three groups: Group I: SC20-resistance gene (current study), *Rsc-ps* (Yang et al. 2013); Group II: *Rsv1* candidate 3Gg2 (Hayes et al. 2004), *Rsc3Q* and *Rsc-pm* (Zheng et al. 2014) and (Yang et al. 2013); and Group III: *Rsc11* (Bai et al. 2009), *Rsc12* (Ma et al. 2010) and *Rsc14Q* (Ma et al. 2011) based on their order and genetic distance to markers Sct_033 (a), BARC-

SOYSSR_13_1185 (b), BARCSOYSSR_13_1155(c), BARCSOYSSR_13_1140 (d), BARCSOYSSR_13_1136 (e) and BARCSOYSSR_13_1114 (f), Satt334/BARCSOYSSR_13_1099 (g) and SOYHSP176 (h). **b** The gene/loci of the *Rsc-ps* and SC20-resistance genes; *Rsc-ps* gene was flanked by BARCSOYSSR_13_1140 and BARCSOYSSR_13_1155 (380.8-kb region); SC20-resistance gene was flanked by gm-ssr_13-14–gm-indel_13-3 (79-kb region)

Amino acid variation causing resistance/susceptibility to SMV

From the qRT-PCR analysis, we identified two candidate genes (*Glyma.13G194700* and *Glyma.13G195100*) having functional implication after SC20 infection (Fig. 3). To further understand what causes resistance or susceptibility in parental lines, we investigated the amino acid patterns of the two candidates *Glyma.13G194700* and *Glyma.13G195100*. When the two genes' coding sequences in parental lines are compared, the SNPs and Indels between them were found (Fig. 4). Totally, 63 SNPs and 3 Indels were found in *Glyma.13G194700* and 17 SNPs and one Indel found in *Glyma.13G195100*. These variations resulted changes in *Glyma.13G194700* and *Glyma.13G195100* amino acid sequences. Additionally, the amino acid changes were confined to the various parts of TIR, NB and LRR domains (Fig. 5). All the domains in NBS–LRR proteins are required to mediate resistance to pathogens in plants, but play distinct roles in the resistance response. It is possible that small amino acid differences may play a key role in resistance. For instance, a susceptibility reaction resulted in a change

in single amino acid in the LRR of the *Pi-ta* protein of the blast-resistance gene in rice (Bryan et al. 2000) and six amino acid changes in the LRR of the P- and P2-resistance proteins of flax determine specificity differences (Dodds et al. 2001). Considering the amino acid variation, we conclude that two parental lines have different alleles and the allelic variations exist between Qihuang-1 and NN1138-2. These changes could then account for the resistance to SMV in Qihuang-1 and for the susceptibility to SMV in NN1138-2.

Comparison of SC20-resistance genes with other SMV-resistance genes on Chromosome 13

The SC20-resistance genes, *Glyma.13G194700* and *Glyma.13G195100*, that we discovered in this study were located on Chromosome 13 (LG F), while *Rsv1* locus conditioning the resistance to G1–G6 strains from US was the first reported SMV-resistance gene on Chromosome 13 (Yu et al. 1994; Li et al. 2006; Hayes et al. 2004). Previous studies showed that, *Rsv1* is a compilation of several *R* genes close to each other or several alleles of the same gene (Gore et al. 2002; Ma et al. 2003; Li et al. 2006; Yang et al. 2013). *Rsv1* locus is the most complicated one, though the knowledge

of this locus is still limited. However, *Rsv1* is located in a genomic region rich in *R* genes conferring resistance to the various pathogens and strains of NCSI SMV strain system from China (Diers et al. 1992; Gore et al. 2002; Ashfield et al. 2004; Ma et al. 2011; Yang et al. 2013; Zheng et al. 2014; Pham et al. 2015). Therefore, studies focusing on this region should be of a broader scope that addresses a series of questions such as: What are the relationships of different resistant genes in this region? How did these genes change over time? And most importantly, how did the diversified pathogen-resistance functions evolve? Answering these questions, undoubtedly, is a long-lasting task and requires knowledge from multiple aspects.

We utilized the available literature knowledge to compare the different SMV-resistance genes located on Chromosome 13. The markers of SOYHSP176, Satt334, BARCSOYSSR_13_1099, BARCSOYSSR_13_1114, BARCSOYSSR_13_1136, BARCSOYSSR_13_1140, BARCSOYSSR_13_1155, BARCSOYSSR_13_1185 and Sct_033 are those linked to several SMV-resistance genes located on Chromosome 13. Among them Satt334, and BARCSOYSSR_13_1099 shared the common genomic region. According to these nine markers and their genomic position, the SMV-resistant genes presently reported on Chromosome 13 can be classified into four groups (Fig. 6a). The first group is placed to an interval between markers BARCSOYSSR_13_1140 and BARCSOYSSR_13_1185 and includes *Rsc-ps* related with resistance to SC7 in PI96983 (Yang et al. 2013) and SC20-resistance genes, *Glyma.13G194700* and *Glyma.13G195100*, from Qihuang-1 in the current study. The second group is located between BARCSOYSSR_13_1114 and BARCSOYSSR_13_1136. The members of this group include the *Rsv1* strong candidate *3Gg2* (Hayes et al. 2004) and *Rsc-pm* from PI 96983 as well as *Rsc3Q* from Qihuang-1 (Yang et al. 2013; Zheng et al. 2014). The third group contains *Rsc11*, *Rsc14Q* and *Rsc12* from Qihuang-1 and Qihuang-22 (Bai et al. 2009; Ma et al. 2010, 2011), which is not well-documented sharing common region with the first two groups, and to clarify the relationship between them and others need an additional fine-mapping. The last group is located to an interval flanked by Satt334/BARCSOYSSR_13_1099 and SOYHSP176 and contains SC18-resistance gene from Qihuang-22 (Li et al. 2015); this group is different from the first three groups and its fine map and precise candidate genes have not been reported yet. To discuss the third and fourth group, which includes the resistance genes *Rsc11*, *Rsc12*, *Rsc14Q* and SC18-resistance gene, the available research information is not sufficient; additional research is required in future.

However, there is no overlapping between the mapping regions of the first group (*Rsc-ps* and SC20-resistance gene-Gm13:30501679–30880254) and second group (*Rsc-pm* and *Rsc3Q*-Gm13: 29,815,195–30,464,870) based on the

locations of flanking markers after delimiting the gene region according to the recently released new assembly of the soybean genome (*Glyma.Wm82.a2*). The *Rsc-ps* gene in cultivar PI96983, which confers resistance to SMV strain SC7 (Yang et al. 2013), was flanked by BARCSOYSSR_13_1140 and BARCSOYSSR_13_1155, two markers that delimit a 380.8-kb region (Gm13: 30,501,679–30,880,254) (Fig. 6b). However, the potential candidate gene(s) to *Rsc-ps* has not been reported. Perhaps the SC20-resistance gene in cultivar Qihuang-1 found in this study, mapped to a 79-kb location (Gm13: 30,795,177 and 30,875,000 bp) was in the *Rsc-ps* genomic region (Fig. 6b). But the relationship between genes from the two sources is unknown. The two closely linked TIR–NB–LRR genes, *Glyma.13G194700* and *Glyma.13G195100* were considered potential ones for SC20 resistance in Qihuang-1. It is possible that SC20-resistance gene(s) and *Rsc-ps* are firmly related (but different genes), or identical as a same gene, or different allele of a same gene, however, this needs further research.

Author contribution statement JG, HZ and AK designed the methods and experiments. AK, CL, JY, NL, YY, YS and RR conducted SMV inoculations and resistance evaluation, and managed field work. AK and CL performed the genotype analysis. AK performed the qRT-PCR and sequencing. AK analyzed the data. AK, KL and JG drafted the manuscript. All authors critically read and approved the final manuscript.

Acknowledgements This work was financially supported through grants from the China National Key R & D Program for crop Breeding (2016YFD0100304, 2017YFD0101500), the Natural Science Foundation of China (31671718, 31571695), the MOE 111 Project (B08025), the MOE Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT_17R55), the MOA CARS-04 program, the Jiangsu Higher Education PAPD Program, the Fundamental Research Funds for the Central Universities (Y0201600115) and the Jiangsu JCIC-MCP. The funders had no role in work design, data collection and analysis, or decision and preparation of the manuscript.

Compliance with ethical standards

Conflict of interest The authors have declared that no competing or conflicts of interest exist.

Ethical standards The experiments were performed in compliance with the current laws of China.

References

- Arif M, Hassan S (2002) Evaluation of resistance in soybean germplasm to Soybean mosaic potyvirus under field conditions. *J Biol Sci* 2:601–604

- Ashfield T, Danzer J, Held D, Clayton K, Keim P, Maroof MS, Webb D, Innes R (1998) Rpg1, a soybean gene effective against races of bacterial blight, maps to a cluster of previously identified disease resistance genes. *Theor Appl Genet* 96:1013–1021
- Ashfield T, Ong LE, Nobuta K, Schneider CM, Innes RW (2004) Convergent evolution of disease resistance gene specificity in two flowering plant families. *Plant Cell* 16:309–318
- Bai L, Li HC, Ma Y, Wang DG, Liu N, Zhi HJ (2009) Inheritance and gene mapping of resistance to soybean mosaic virus strain SC-11 in soybean. *Soybean Sci* 28:1–6
- Bryan GT, Wu KS, Farrall L, Jia Y, Hershey HP, McAdams SA, Faulk KN, Donaldson GK, Tarchini R, Valent B (2000) A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *Plant Cell* 12:2033–2046
- Chen ML, Luo J, Shao GN, Wei XJ, Tang SQ, Sheng ZH, Song J, Hu PS (2012) Fine mapping of a major QTL for flag leaf width in rice, *qFLW4*, which might be caused by alternative splicing of *NALI*. *Plant Cell Rep* 31:863–872
- Cheng HJ, Guo SG, Yu ZL (1982) Major diseases and pests of soybeans in China proceedings of the first China/USA soybean symposium and working group meeting, pp 74–80, 26–30 July 1982
- Diers B, Mansur L, Imsande J, Shoemaker R (1992) Mapping Phytophthora resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Sci* 32:377–383
- Dodds PN, Lawrence GJ, Ellis JG (2001) Six amino acid changes confined to the leucine-rich repeat β -strand/ β -turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. *Plant Cell* 13:163–178
- Doyle JJ (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Dutta S, Kumawat G, Singh BP, Gupta DK, Singh S, Dogra V, Gaikwad K, Sharma TR, Raje RS, Bandhopadhyaya TK (2011) Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus cajan* (L.) Mills paugh]. *BMC Plant Biol* 11:1–13
- Fu SX, Zhan Y, Zhi HJ, Gai JY, Yu DY (2006) Mapping of SMV resistance gene *Rsc7* by SSR markers in soybean. *Genetica* 128:63–69
- Gore MA, Hayes AJ, Jeong SC, Yue YG, Buss GR, Saghai Maroof MA (2002) Mapping tightly linked genes controlling potyvirus infection at the *Rsv1* and *Rpv1* region in soybean. *Genome* 45:592–599
- Guo D, Zhi H, Wang Y, Gai J, Zhou X, Yang C (2005) Identification and distribution of strains of *Soybean mosaic virus* in middle and northern of Huang Huai Region of China. *Soybean Sci* 27:64–68
- Guo DQ, Wang YW, Zhi HJ, Gai JY, Li HC, Li K (2007) Inheritance and gene mapping of resistance to SMV strain group SC-13 in soybean. *Soybean Sci* 26:21–24
- Hayes A, Jeong S, Gore M, Yu Y, Buss G, Tolin S, Maroof MS (2004) Recombination within a nucleotide-binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to soybean mosaic virus in soybeans. *Genetics* 166:493–503
- Hirata K, Masuda R, Tsubokura Y, Yasui T, Yamada T, Takahashi K, Nagaya T, Sayama T, Ishimoto M, Hajika M (2014) Identification of quantitative trait loci associated with boiled seed hardness in soybean. *Breed Sci* 64:362–370
- Jeong S, Maroof S (2004) Detection and genotyping of SNPs tightly linked to two disease resistance loci, *Rsv1* and *Rsv3*, of soybean. *Plant Breeding* 123:305–310
- Jeong SC, Kristipati S, Hayes AJ, Maughan PJ, Noffsinger SL, Gunduz I, Bussa GR, Saghai Maroof MA (2002) Genetic and sequence analysis of markers tightly linked to the *Soybean mosaic virus* resistance gene, *Rsv3*. *Crop Sci* 42:265–270
- Jiang S, Zhang X, Zhang F, Jiang S, Zhang X, Zhang F, Xu Z, Chen W, Li Y (2012) Identification and fine mapping of qCTH4, a quantitative trait loci controlling the chlorophyll content from tillering to heading in rice (*Oryza sativa* L.). *J Hered* 103:720–726
- Kosambi DD (1943) The estimation of map distances from recombination values. *Ann Eugen* 12:172–175
- Li HC, Zhi HJ, Gai JY, Guo DQ, Wang YW, Li K, Bai L, Yang H (2006) Inheritance and gene mapping of resistance to soybean mosaic virus strain SC14 in soybean. *J Integr Plant Biol* 48:1466–1472
- Li K, Ren R, Karthikeyan A, Gao L, Yuan Y, Liu ZT, Zhong YK, Zhi HJ (2015) Genetic analysis and identification of two soybean mosaic virus resistance genes in soybean [*Glycine max* (L.) Merr.]. *Plant Breed* 134(6):684–695
- Li K, Yang QH, Zhi HH, Gai JY (2010) Identification and distribution of soybean mosaic virus strains in southern China. *Plant Dis* 94:351–357
- Liu W, Li X, Zhou K, Pan X, Li Y, Lu T, Sheng X (2016) Mapping of QTLs controlling grain shape and populations construction derived from related residual heterozygous lines in rice. *J Agric Sci* 8:104–114
- Ma G, Chen PY, Buss G, Tolin SA (2003) Genetic study of a lethal necrosis to soybean mosaic virus in PI 507389 soybean. *J Hered* 94:205–211
- Ma Y, Li HC, Wang DG, Liu N, Zhi HJ (2010) Molecular mapping and marker assisted selection of soybean mosaic virus resistance gene *RSC12* in soybean. *Legume Genom Genet* 1(8):1–6
- Ma Y, Wang DG, Li HC, Zheng GJ, Yang YQ, Li HW, Zhi HJ (2011) Fine mapping of the *RSC14Q* locus for resistance to soybean mosaic virus in soybean. *Euphytica* 181:127–135
- Maroof M, Tucker DM, Skoneczka JA, Bowman BC, Tripathy S, Tolin SA (2010) Fine mapping and candidate gene discovery of the soybean mosaic virus resistance gene, *Rsv4*. *Plant Genome* 3:14–22
- Michelmore RW, Paran I, Kesseli R (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci* 88:9828–9832
- Mohan M, Nair S, Bhagwat A, Krishna T, Yano M, Bhatia C, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol Breed* 3:87–103
- Moon JK, Jeong SC, Van K, Maroof MS, Lee SH (2009) Marker-assisted identification of resistance genes to soybean mosaic virus in soybean lines. *Euphytica* 169:375–385
- Pham AT, Harris DK, Buck J, Hoskins A, Serrano J, Abdel-Haleem H, Cregan P, Song QJ, Boerma HR, Li ZL (2015) Fine mapping and characterization of candidate genes that control resistance to *Cercospora sojina* K. Hara in two soybean germplasm accessions. *PLoS ONE* 10:e0126753
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song QJ, Thelen JJ, Cheng JL, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu SQ, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du JC, Tian ZX, Zhu LC, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinzaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
- Shirasawa K, Isobe S, Tabata S, Hirakawa H (2014) Kazusa Marker DataBase: a database for genomics, genetics, and molecular breeding in plants. *Breed Sci* 64:264–271
- Song QJ, Marek L, Shoemaker R, Lark K, Concibido V, Delannay X, Specht JE, Cregan P (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet* 109:122–128
- Song QJ, Jia GF, Zhu YL, Grant D, Nelson RT, Hwang EY, Hyten DL, Cregan P (2010) Abundance of SSR motifs and development of candidate polymorphic SSR markers (BARCSOYSSR_1.0) in soybean. *Crop Sci* 50:1950–1960
- Song J, Liu Z, Hong H, Ma Y, Tian L, Li X, Li Y-H, Guan R, Guo Y, Qiu L-J (2016) Identification and validation of loci governing seed

- coat color by combining association mapping and bulk segregation analysis in soybean. *PLoS ONE* 11(7):e0159064
- Su CF, Lu WG, Zhao TJ, Gai JY (2010) Verification and fine-mapping of QTLs conferring days to flowering in soybean using residual heterozygous lines. *Chin Sci Bull* 55:499–508
- Suh SJ, Bowman BC, Jeong N, Yang K, Kastl C, Tolin SA, Maroof M, Jeong SC (2011) The *Rsv3* locus conferring resistance to soybean mosaic virus is associated with a cluster of coiled-coil nucleotide-binding leucine-rich repeat genes. *Plant Genome* 4:55–64
- Tang S, Shao G, Wei X, Chen M, Sheng Z, Luo J, Jiao G, Xie L, Hu Tang S, Shao G, Wei X, Chen M, Sheng Z, Luo J, Jiao G, Xie L, Hu P (2013) QTL mapping of grain weight in rice and the validation of the QTL *qTGW3.2*. *Gene* 527:201–206
- Tollenaere R, Hayward A, Dalton-Morgan J, Campbell E, Lee JR, Lorenc MT, Manoli S, Stiller J, Raman R, Raman H (2012) Identification and characterization of candidate Rlm4 blackleg resistance genes in *Brassica napus* using next-generation sequencing. *Plant Biotechnol J* 10:709–715
- Tuyen D, Lal S, Xu DH (2010) Identification of a major QTL allele from wild soybean (*Glycine soja* Sieb. & Zucc.) for increasing alkaline salt tolerance in soybean. *Theor Appl Genet* 121:229–236
- Van Ooijen J (2006) JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*
- Wang XQ, Gai JY, Pu ZQ (2003) Classification and distribution of strains of Soybean mosaic virus middle and lower Huang-Huai and Changjiang Valleys. *Soybean Sci* 22:102–107
- Wang YJ, Dong Fang Y, Wang XQ, Yang YL, Yu DY, Gai JY, Wu XL, He CY, Zhang JS, Chen SY (2004) Mapping of five genes resistant to SMV strains in soybean. *Acta Genet Sin* 31:87–90
- Wang YW, Zhi HJ, Guo DQ, Gai JY, Chen QS, Li K, Li HC (2005) Classification and distribution of strains of Soybean mosaic virus in northern china spring planting soybean region. *Soybean Sci* 24:263–268
- Wang DG, Ma Y, Liu N, Yang ZL, Zheng GJ, Zhi HJ (2011a) Fine mapping of resistance to soybean mosaic virus strain SC4 in soybean based on genomic-SSR markers. *Plant Breed* 130:653–659
- Wang DG, Ma Y, Yang YQ, Liu N, Li CY, Song YP, Zhi HJ (2011b) Fine mapping and analyses of *RSC8* resistance candidate genes to soybean mosaic virus in soybean. *Theor Appl Genet* 122:555–565
- Wang DG, Li HW, Zhi HJ, Tian Z, HuC, Hu GY, Huang ZP, Zhang L (2014) Identification of strains and screening of resistance resources to soybean mosaic virus in Anhui Province. *Chin J Oil Crop Sci* 36:374–379
- Wang YF, Lu JJ, Chen S, Shu L, Palmer RG, Xing GN, Li Y, Yang SP, Yu DY, Zhao TJ, Gai JY (2014b) Exploration of presence/absence variation and corresponding polymorphic markers in soybean genome. *J Integr Plant Biol* 56:1009–1019
- Xu R, Shi CE, Zhang LF (2004) Utilization of Qihuang 1 in soybean breeding in the Huanghuai region. *J Plant Genet Resour* 5:170–175
- Yamanaka N, Watanabe S, Toda K, Hayashi M, Fuchigami H, Takahashi R, Harada K (2005) Fine mapping of the *FT1* locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. *Theor Appl Genet* 110:634–639
- Yamanaka N, Fuentes F, Gilli J, Watanabe S, Harada K, Ban T, Abdelnoor R, Nepomuceno A, Homma Y (2006) Identification of quantitative trait loci for resistance against soybean sudden death syndrome caused by *Fusarium tucumaniae*. *Pesqui Agropecu Bras* 41:1385–1391
- Yan H, Wang H, Cheng H, Hu Z, Chu S, Zhang G, Yu D (2015) Detection and fine-mapping of SC7 resistance genes via linkage and association analysis in soybean. *J Integr Plant Biol* 57:722–729
- Yang QH, Gai JY (2010) Identification, inheritance and gene mapping of resistance to a virulent soybean mosaic virus strain SC15 in soybean. *Plant Breed* 130:128–132
- Yang YQ, Zheng GJ, Han L, Wang DG, Yang XF, Yuan Y, Huang SH, Zhi HJ (2013) Genetic analysis and mapping of genes for resistance to multiple strains of Soybean mosaic virus in a single resistant soybean accession PI 96983. *Theor Appl Genet* 126:1783–1791
- You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD (2008) BatchPrimer3: a high through put web application for PCR and sequencing primer design. *BMC Bioinform* 9:1
- Yu YG, Saghai MMA, Buss GR, Maughan PJ, Tolin SA (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. *Phytopathology* 84:60–64
- Zhan Y, Zhi HJ, Yu DY, Gai JY (2006) Identification and distribution of SMV strains in Huang-Huai Valleys. *Sci Agric Sin* 39:2009–2015
- Zheng GJ, Yang YQ, Ma Y, Yang XF, Chen SY, Ren R, Wang DG, Yang ZL, Zhi HJ (2014) Fine mapping and candidate gene analysis of resistance gene *RSC3Q* to soybean mosaic virus in Qihuang No. 1. *J Integr Agric* 13:2608–2615