Theor Appl Genet (2011) 122:445–457
DOI 10.1007/s00122-010-1459-6

ORIGINAL PAPER

Multiplex single nucleotide polymorphism (SNP) assay for detection of soybean mosaic virus resistance genes in soybean

Ainong Shi · Pengyin Chen · Richard Vierling · Cuming Zheng · Dexiao Li · Dekun Dong · Ehsan Shakiba · Innan Cervantez

Received: 15 March 2010 / Accepted: 13 September 2010 / Published online: 8 October 2010

Abstract *Soybean mosaic virus* (SMV) is one of the most destructive viral diseases in soybean (*Glycine max*). Three independent loci for SMV resistance have been identified in soybean germplasm. The use of genetic resistance is the most effective method of controlling this disease. Marker assisted selection (MAS) has become very important and useful in the effort of selecting genes for SMV resistance. Single nucleotide polymorphism (SNP), because of its abundance and high-throughput potential, is a powerful tool in genome mapping, association studies, diversity analysis, and tagging of important genes in plant genomics. In this study, a 10 SNPs plus one insert/deletion (InDel) multiplex assay was developed for SMV resistance: two SNPs were developed from the candidate gene *3gG2* at *Rsv1* locus, two SNPs selected from the clone N11PF linked to *Rsv1*, one 'BARC' SNP screened from soybean chromosome 13 [linkage group (LG) F] near *Rsv1*, two 'BARC' SNPs from probe A519 linked to *Rsv3*, one 'BARC' SNP from chromosome 14 (LG B2) near *Rsv3*, and two 'BARC' SNPs from chromosome 2 (LG D1b) near *Rsv4*, plus one InDel marker from expressed sequence tag (EST) AW307114 linked to *Rsv4*. This 11 SNP/InDel multiplex assay showed polymorphism among 47 diverse soybean germplasm, indicating this assay can be used to investigate the mode of inheritance in a SMV resistant soybean line carrying *Rsv1*, *Rsv3*, and/or *Rsv4* through a segregating population with phenotypic data, and to select a specific gene or pyramid two or three genes for SMV resistance through MAS in soybean breeding program. The presence of two SMV resistance genes (*Rsv1* and *Rsv3*) in J05 soybean was confirmed by the SNP assay.

Communicated by I. Rajcan.

A. Shi
Syngenta Seeds Inc., 2369 330th St., Slater, IA 50224, USA

P. Chen (✉) · D. Dong · E. Shakiba · I. Cervantez
Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR 72701, USA
e-mail: pchen@uark.edu

A. Shi · R. Vierling
Indiana Crop Improvement Association and Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA

C. Zheng
Syngenta Seeds Inc., 14031 Trestle Road, Highland, IL 62249, USA

D. Li
College of Agriculture, Northwest University of Agriculture and Forestry, Yangling, Shaanxi 712100, People's Republic of China

Introduction

Soybean mosaic virus (SMV) is one of the most prevalent and destructive viral diseases in soybean [*Glycine max* (L.) Merr.]. The use of genetic resistance is the most effective method of controlling this disease. Three independent loci, *Rsv1*, *Rsv3*, and *Rsv4* conferring SMV resistance have been identified and mapped on different chromosomes (linkage groups (LG)) of soybean (Hayes et al. 2000; Jeong et al. 2002; Yu et al. 1994). The *Rsv1* locus was the first resistance gene identified and is the most common in soybean germplasm. Nine alleles have been reported at the *Rsv1* locus: *Rsv1* in PI 96983, *Rsv1-h* in 'Suweon 97', *Rsv1-k* in 'Kwanggyo', *Rsv1-m* in 'Marshall', *Rsv1-n* in PI 507389, *Rsv1-r* in 'Raiden', *Rsv1-s* in LR1, *Rsv1-t* in 'Ogden', and *Rsv1-y* in 'York' (Chen and Choi 2008). These *Rsv1* alleles usually confer resistance to less virulent SMV strains, but

susceptible or necrotic reactions to more virulent strain. The *Rsv3* locus confers resistance to more virulent strain, but susceptibility to less virulent strain. The *Rsv4* locus confers resistance to all SMV strains at early stage, but middle susceptibility at late stage.

Rsv1 was mapped on soybean chromosome 13 (LG F) (Yu et al. 1994). Two Restriction Fragment Length Polymorphism (RFLP) markers, pA 186 and pK 644a, and one Simple Sequence Repeat (SSR) marker, SM176, were found to be tightly linked to *Rsv1* with distances of 1.5, 2.1, and 0.5 cM, respectively (Yu et al. 1994). A Random Amplification of Polymorphic DNA (RAPD) marker OPN-11_{1000} was also described as a linked marker to *Rsv1* by Li et al. (1998). Subsequently, the OPN-11 was mapped 1 cM away from *Rsv1* (Gore et al. 2002) and its derived Sequence Characterized Amplified Region (SCAR) marker was reported to be closely linked (1.8 ± 1.3 cM) in coupling phase with the *Rsv1* locus (Li et al. 2004). Later, the PCR fragment amplified from the SCAR primer pair was named N11PF and eight fragments amplified from eight different soybean genotypes using the SCAR primer pair were sequenced and the sequences were deposited in the GenBank (accession no. AY305014–AY305021) (Jeong and Saghai Maroof 2004). Ten allele-derived SNPs and seven Insert/Deletion (InDel) polymorphisms were found in the eight N11PF fragments and four of them were confirmed to be polymorphic by AS-Allele-Specific Polymerase Chain Reaction (AS-PCR) in parents and two F_2 populations derived from PI 96983 × Lee 68 and 'Myeongjunamulkong' × Lee 68, respectively (Jeong and Saghai Maroof 2004). Another two SNP markers were developed from L20a (GenBank accession no. AF322632), a Toll and Interleukin-1 Receptor (TIR)-Nucleotide Binding Site (NBS) genomic sequence that was mapped close to *Rsv1* (Hayes and Saghai Maroof 2000) and confirmed to be polymorphic by AS-PCR (Jeong and Saghai Maroof 2004). Four Amplified Fragment Length Polymorphism (AFLP) derived dominant markers on NBS region, R11 (518 bp), R12 (171 bp), R13 (261 bp), and R14 (312 bp), were identified to be linked to *Rsv1* in the F_2 population derived from PI96983 (*Rsv1*) × Lee68 (*rsv*) by bulk segregant analysis, and the four markers were mapped by RFLP to chromosome 13 within 3.5 cM to *Rsv1* and the closest one R11 was within the gene less than 0.4 cM (Hayes and Saghai Maroof 2000). In another study, a high-resolution map of the *Rsv1* region containing 38 loci was constructed with 24 markers including one RAPD (OPN11), four SSR (HSP176, 64-A8C, Satt510, and Satt120), and 19 RFLP markers; *Rsv1* was flanked by HSP176 (2.9 cM) and Satt510 (2.4 cM); and the closest SSR marker to *Rsv1* was 64-A8C (0.5 cM) (Gore et al. 2002). The *Rsv1* gene in J05 soybean was mapped on chromosome 13 (LG F) and flanked by Sat_154 and Satt510 with 0.5 and 2.3 cM, respectively (Shi et al. 2008a). The RAPD marker $OPN11_{980/1070}$ and its derived SCAR marker $SCN11_{980/1070}$ were also identified to be linked to *Rsv1* with a distance of 3.03 cM in another research (Zheng et al. 2003).

Rsv1 locus is actually a multi-gene family in which closely linked members contribute to the host-pathogen interaction (Hayes et al. 2004). Twelve NBS resistance gene analogs were mapped at the *Rsv1* locus in PI 96983 (Jeong et al. 2001). Additionally, six clones for the region around the *Rsv1* locus on chromosome 13 (LG F) associated with SMV resistance were reported (Hayes et al. 2004) and three of them were completely sequenced (GenBank accession no. AY518517–AY518519). One of the six clones, *3gG2* gene, co-segregated with *Rsv1* and the distance between *3gG2* and *Rsv1* was 0 cM in a population derived from PI 96983. 'Marshall' carrying *Rsv1-m* and 'Ogden' carrying *Rsv1-t* (Chen et al. 2001) were also reported to contain the *3gG2* gene (Hayes et al. 2004). Recently, a PCR-based primer Rsv1-f/r was developed from the *Rsv1* candidate gene *3gG2* and the study confirmed presence of *3gG2* gene in all genotypes tested that carry all *Rsv1* alleles (*Rsv1, Rsv1-h, Rsv1-k, Rsv1-m, Rsv1-n, Rsv1*-r, and *Rsv1-t*) except *Rsv1-y* (*Rsv1-s* was not been tested) (Shi et al. 2008b). In this study, the SSR marker Satt114 was found to be linked to Rsv1-f/r with a distance of 5.42 cM. *Rsv3* was mapped on chromosome 14 (LG B2) (Jeong et al. 2002). The RFLP marker A519 and its allele-derived PCR-based marker, A519F/R, were identified to be linked to *Rsv3* with a distance of 0.9 cM based on the data from 195 $F_{2:3}$ lines derived from L29 × Lee 68 and the SSR marker Satt63 was also mapped at same side of A519 with a distance of 6.5 cM to *Rsv3* (Jeong et al. 2002). Another allele-derived SSR marker M3Satt was mapped at another side of *Rsv3* with a 0.8 cM distance in the population of Tousan 140 × Lee 68 (Jeong et al. 2002). *Rsv3* in J05 was also mapped on chromosome 14 (LG B2) with a distance of 1.5 cM from Sat_424 and 2.0 cM from Satt726 (Shi et al. 2008a). Three SNP markers were developed from the sequence of the A519 allele and confirmed by AS-PCR between Lee 68 and L29 (Jeong and Saghai Maroof 2004). *Rsv4* in V94-5152 was mapped on chromosome 2 (LG D1b) and flanked by two SSR markers, Satt542 at 4.7 cM and Satt558 at 7.8 cM (Hayes et al. 2000). Three ESTs (expressed sequence tags) derived PCR-based markers AI 856415-g, AI 856415-S and BF 070293-S were mapped at 2.8 cM on one side of *Rsv4*; additional two EST markers AW 307114A (3.3 cM) and AW 471852R (2.4 cM) were mapped on the other side of *Rsv4* and the SSR marker Satt634 was also mapped on the side of AI 856415-g/-S with a distance 2.2 cM to *Rsv4* on Chromosome 2 (LG D1b) based on the data from the F_2 population of 'Sowonkong' × V94-5152 (Hwang et al. 2006). Saghai Maroof et al. (2010) identified six new SSR-based markers designed from the region of

interest around *Rsv4* of the whole genome shotgun sequence (WGS) of soybean and mapped them in the D26 × Lee 68 and V94-5152 × Lee 68 populations; the closest flanking markers, 212MAT16 and 212MTATA11 shortened the genetic interval containing *Rsv4* to 0.7 cM and 1.3 cM, respectively; and the physical distance between these two markers based on the available Williams82 sequence was approximately 100 kb.

Marker assisted selection (MAS) has been successfully used in selecting for specific genes/alleles or combination of multiple resistance genes/alleles for disease resistance in plants by using genetic markers that are linked to the resistance genes (Brahm et al. 2000; Fjellstrom et al. 2004; Foolad and Sharma 2005). Single nucleotide polymorphism (SNP), with its abundance, cost efficiency and high-throughput scoring potential, has become a powerful tool in genome mapping, association studies, diversity analysis, and tagging of important genes in plant genomics (Rafalski 2002; Giancola et al. 2006; Caicedo et al. 2007; Choi et al. 2007). SNPs have been discovered and a SNP genetic map has been constructed in soybean (Choi et al. 2007; http://bfgl.anri.barc.usda.gov/soybean/index.html). SNP markers have been successfully used in detection of the *Rcs3* gene for resistance to frogeye leaf spot and *Rpp3* locus for resistance to Asian soybean rust in soybean (Missaoui et al. 2007; Hyten et al. 2007). The objective of this research is to develop multiplex SNP/InDel assays for identification and selection of SMV resistance alleles through MAS in soybean breeding program.

Materials and methods

Plant materials and SMV tests

Forty-seven diverse soybean genotypes were selected for this research (Table 1). The reactions of the selected genotypes to SMV strains G1 to G7 (Cho and Goodman 1979) were evaluated in the greenhouse. Most of soybean genotypes in this research were previously screened with SMV strains G1 and G7, and resistance alleles have been proposed (Zheng et al. 2005). In addition, some of these genotypes were later tested with five or seven SMV strains in another experiment (Zheng et al. 2008). Seeds from each of the 47 soybean genotype were planted in 15-cm-diameter plastic pots each containing with six plants in the greenhouse. Each genotype was inoculated with seven SMV strains (G1 to G7).

A cross between J05 (*Rsv1 Rsv3*) and a susceptible cultivar ''Essex'' (*rsv*) was made for validation of SNP markers linked to *Rsv1* and *Rsv3*. J05 was shown to contain two SMV resistance genes, *Rsv1* and *Rsv3*, based on classic genetic analysis and allelism tests (Zheng et al. 2006) and later confirmed by linkage analysis using SSR markers (Shi et al. 2008a). F_1 plants were grown without SMV inoculation to generate an F_2 population. The reactions of F_2 plants to SMV G1 (referred to as G1 hereafter) were used as phenotypic data for the presence of *Rsv1* because *Rsv1* confers resistance to G1 and *Rsv3* gives rise to susceptibility to G1. The plants susceptible to G1 from the F_2 population were harvested separately to develop $F_{2:3}$ lines. The reactions to SMV G7 (referred to as G7 hereafter) of those $F_{2:3}$ lines derived from the G1-susceptible F_2 plants were used as phenotypic data for the presence of *Rsv3* because these $F_{2:3}$ lines did not contain *Rsv1* and therefore, inoculation with G7 would reveal the presence or absence of *Rsv3*.

The F_2 population of J05 × Essex was growing in the pots at greenhouse and all the single F_2 plants were inoculated with G1. Subsequently, the $F_{2:3}$ lines derived from the G1-susceptible F_2 plants of the cross 'J05' × 'Essex' were inoculated with G7. Approximately 30 plants per $F_{2:3}$ line were used in the G7 inoculation. The greenhouse conditions were maintained at 20–25°C temperature and 15 h photoperiod at the Harry R. Rosen Alternative Pest Control Center of University of Arkansas, Fayetteville, AR. Virus inoculum was prepared by grinding systemically infected leaves from virus-maintaining stock plants with a mortar and pestle in 0.05 M potassium phosphate buffer at an approximate dilution of 1:10 (w/v) at pH 7.2. Inoculation was performed by rubbing the inoculum with a pestle onto both unifoliolate leaves that had been previously dusted with carborundum (Chen et al. 1991; Zheng et al. 2005). SMV strains were kindly provided by Dr. Sue Tolin of Virginia Polytechnic Institute and State University. Plants of each genotype were monitored for SMV symptoms on a regular basis and were classified as resistant (R, symptomless), necrotic (N, stem-tip necrosis), or susceptible (S, mosaic). The $F_{2:3}$ lines from J05 × Essex were classified as homozygous resistant, segregating, or homozygous susceptible based on individual plant reactions.

DNA extraction, PCR amplification and DNA sequencing

Genomic DNA was extracted from fresh leaves of greenhouse-grown plants using the CTAB (hexadecyltrimethyl ammonium bromide) method (Kisha et al. 1997) or from seeds using SDS method with minor modification (Kamiya and Kiguchi 2003). In the SDS procedure for DNA extraction, about 0.1 mg soybean seed powder was mixed with 1.5 ml lysis buffer (200 mM Tris pH 7.5, 25 mM EDTA pH 8.0, 0.5%SDS, 288 mM NaCl) in a 2 ml microcentrifuge tube by vortexing. The mixture was incubated at room temperature for 30 min, and then centrifuged at 10,000 rpm for 10 min. Then, 900 µl

Table 1 Polymorphism of 11 Single nucleotide polymorphism (SNP) markers in 47 diverse soybean genotypes

Genotype	Possible gene	Linked to *Rsv1*					Linked to *Rsv3*			Linked to *Rsv4*		
		3gG2-snp1	3gG2-snp2	N11PF-snp2	N11PF-snp3	Barc_015435-01966	A519-Snp2	A519-Snp4	Barc-012953-00413	Barc-011147-00855	Barc-025955-05182	AW307114-indel1
Lee68	*rsv*	0 0	C C	G G	T T	A A	C C	C C	C T	A A	A A	d d
William	*rsv*	0 0	C C	G G	T T	A A	C C	C C	C T	G G	G G	d d
Camp	*rsv*	0 0[a]	C C	0 0	T T	A A	C C	C C	C T	A A	A A	d d
Essex	*rsv*	G T	A A	T T	T T	C C	C C	C C	C T	A A	A A	d d
PI 96983	*Rsv1*	T T	C C	0 0	T T	C C	G G	T T	C T	A A	G G	A A
V94-3971	*Rsv1*	T T	C C	0 0	T T	C C	C C	C C	C T	A A	A A	d d
Hourei	*Rsv1+3*	T T	C C	G G	G G	C C	G G	T T	C T	G G	G G	A A
J05	*Rsv1+3*	T T	C C	T T	T T	C C	G G	T T	T T	A A	A A	A A
PI 486355	*Rsv1+4*	T T	C C	G G	G G	C C	G G	T T	T T	G G	A A	A A
Kefeng no. 1	*Rsv1-h*	T T	C C	T T	T T	A A	G G	T T	T T	G G	A A	A A
Suweon 97	*Rsv1-h*	T T	C C	G G	G G	C C	C C	C C	C T	A A	G G	d d
'Iwate Wasw Karome' (Bunri)	*Rsv1-k*	T T	C C	G G	G G	C C	C C	C C	C T	A A	G G	d d
Iwate no. 1	*Rsv1-k*	T T	C C	G G	G G	C C	G G	T T	C T	A A	G G	d d
Kwanggyo	*Rsv1-k*	T T	C C	G G	G G	C C	G G	T T	C T	G G	G G	A A
Nohrin no. 3	*Rsv1-k*	T T	C C	G G	G G	C C	G G	T T	C T	G G	G G	d d
Nooki no. 1	*Rsv1-k*	T T	C C	0 0	T T	A A	G G	T T	T T	G G	A A	A A
Shakkin-nashi	*Rsv1-k*	T T	C C	G G	G G	C C	G G	T T	C T	G G	A A	A A
Tohoku no.1	*Rsv1-k*	T T	C C	G G	G G	C C	G G	T T	C T	G G	A A	d d
Marshall	*Rsv1-m*	T T	C C	T T	T T	A A	C C	0 0	C T	G G	G G	A A
PI 61944	*Rsv1-n*	0 0	C C	0 0	T T	C C	G G	T T	T T	G G	G G	d d
PI 61947	*Rsv1-n*	0 0	C C	0 0	T T	C C	G G	T T	T T	G G	G G	d d
V262	*Rsv1-n*	T T	C C	G G	G G	C C	C C	C C	C T	A A	A A	d d
Houjaku Kuwazu	*Rsv1-r*	T T	C C	G G	G G	C C	G G	T T	C T	A A	G G	d d
Miyagi Shirome	*Rsv1-r*	T T	C C	G G	G G	C C	G G	T T	C T	G G	A A	d d
Raiden	*Rsv1-r*	T T	C C	G G	G G	C C	C C	C C	C T	A A	G G	d d
Tsuronoko	*Rsv1-r*	T T	C C	G G	G G	C C	G G	T T	C T	G G	A A	d d
Ogden	*Rsv1-t*	T T	C C	T T	T T	A A	C C	C C	C T	A A	G G	A A
Davis	*Rsv1-y*	G G	C C	T T	T T	A A	C C	C C	C T	A A	A G	A A
York	*Rsv1-y*	G G	C C	T T	T T	A A	C C	C C	C T	G G	A A	d d
Zho 18	*Rsv1-y+3*	G G	C C	T T	T T	C C	G G	T T	C T	G G	A A	d d
Cordell	*Rsv3*	0 0	C C	0 0	T T	A A	G G	T T	T T	A A	G G	A A[b]
Harosoy	*Rsv3*	0 0	C C	G G	T T	A A	G G	T T	T T	G G	A A	A A
V229	*Rsv3*	G T	A A	T T	T T	C C	G G	T T	T T	A A	A A	d d

Table 1 continued

Genotype	Possible gene	Linked to *Rsv1*					Linked to *Rsv3*			Linked to *Rsv4*		
		3gG2-snp1	3gG2-snp2	N11PF-snp2	N11PF-snp3	Barc_015435-01966	A519-Snp2	A519-Snp4	Barc-012953-00413	Barc-011147-00855	Barc-025955-05182	AW307114-indel1
L29	*Rsv3*	G T	C C	G G	T T	A A	0 0	0 0	T T	G G	G G	d d
Da bai ma	*Rsv3+4*	G T	A A	T T	T T	C C	G G	T T	T T	A A	A A	d d
Columbia	*Rsv3+4*	G T	C C	0 0	T T	A A	0 0	C T	0 0	G G	A A	d d
PI 398593	*Rsv4*	0 0	A C	0 0	T T	A A	C C	C C	C T	G G	G G	d d
Kaeri-gnt 220-7	*Rsv4*	G G	C C	T T	T T	A A	G G	T T	T T	G G	A A	A A
V94-5152	*Rsv4*	G T	A A	T T	T T	C C	G G	T T	T T	A A	A A	A A
V94-9003	*Rsv4*	G T	A A	0 0	T T	C C	C C	C C	C T	A A	A A	A A
Viginia	*Rsv4*	G T	A A	G G	T T	A A	G G	T T	T T	G G	A A	d d
William 79	*?*	0 0	C C	G G	T T	A A	C C	C C	C T	G G	G G	d d
William 82	*?*	0 0	C C	G G	T T	A A	C C	C C	C T	G G	G G	d d
Jindou1	*?*	G G	C C	T T	T T	A A	G G	T T	C T	G G	A A	d d
Jiunong 21	*?*	G T	C C	0 0	T T	A A	G G	T T	C T	A A	G G	d d
Ching Tao 21	*?*	T T	C C	G G	G G	C C	G G	T T	C T	G G	A A	d d
Suzuyataka	*?*	T T	C C	G G	G G	C C	G G	T T	T T	G G	A A	A A

[a] The SNP type such as [G T] signifies a heterogeneous SNP type, [T T] and [G G] signify a homogenous SNP type, [0 0] signifies no data, dd signifies deletion marker

[b] The AW307114-indel1 is an Insert/deletion marker, [A A] signifies an insert and [d d] a deletion

Table 2 Six SNP (In/Del) markers detected from the N11PF allele among 5–8 sequences from GenBank

GenBank accession	Genotype	N11PF-snp1	N11PF-snp2	N11PF-snp3	N11PF-snp4	N11PF-In/Del1	N11PF-InDel-snp5
AY305014	Hwangkeumkong	47A[a]	635G	929G	1015T	220–273 Insert1[b]	227T
AY305021	Taekwangkong	47A	635G	929G	1015T	220–273 Insert1	227T
AY305017	Myeungjunamulkong	47A	635G	929G	1015T	220–273 Insert1	227T
AY305018	Peking	47G	638T	923T	1009C	220–273 Insert2	227G
AY305019	PI96983	47G	631T	916T	1002C	220–273 Insert2	227G
AY305016	Lee68	47A	581G	827T	913T	–	–
AY305015	IT182932	47A	581G	827T	913T	–	–
AY305020	Sowon	47A	580T	826T	912T	–	–

[a] The SNP location and type, e.g. 47A signifies the SNP N11PF-snp1 has an 'A' base at location of 47 site of the AY305014 sequence

[b] '220–273 Insert1' represents the DNA sequence (ACGAATATTGTCTACCACTTCCAGCACAATTACAGAAATGTCCATAAAATATCA) and '220–273 Insert2' represents another DNA sequence (ACGAATAGTGTCTACCACTTCCAGCACAATTACAGAAATGTCCATAAAA-TATCA) located at 220–273 bases of corresponding GenBank accessions; '–" represents the absent of the DNA sequence in the corresponding location. One [T/G] SNP N11PF-InDel-snp5 was also observed in the two insert sequences: '220–273 Insert1' and '220–273 Insert2'

supernatant was transferred to a new 2 ml tube and mixed with 300 µl 5 M potassium acetate by gently inverting 6-8 times. After centrifuged at 10,000 rpm for 10 min, 1 ml supernatant was transferred to another new 2-ml tube and 900 µl isopropanol was added to precipitate the DNA. After mixing well, the mixture was centrifuged at 10,000 rpm for 10 min. The supernatant was discarded and the DNA pellet was dried at room temperature overnight. The DNA pellet was then dissolved in 0.2 ml of 0.1× TE buffer (10 mm Tris–HCl, 1 mM EDTA; pH 8.0). DNA concentration was measured using NanoDrop ND-1000 UV–Vis Spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE).

PCR amplification was performed in an iCycler Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA) or an Eppendorf thermal cycler (Eppendorf, Westbury, NY) following standard PCR procedures with minor modifications. Briefly, each 50 µl PCR reaction mixture consisted of 29.8 µl sterilized ddH_2O, 10 µl 5× Mango Taq reaction buffer (Bioline, London, UK), 3 µl $MgCl_2$ (25 mM), 1.5 µl dNTP (2.5 mM each), 1.5 µl each primer (5 µM), 0.2 µl Mango Taq DNA polymerase (5 U/µl) (Bioline, London, UK), and 2.5 µl template DNA (30 ng/µl). PCR procedure consisted of an initial denaturation step at 95°C for 1 min, 40 cycles of 30 s at 94°C, 30 s at 54°C, and 1 min at 72°C followed by an extension step at 72°C for 5 min and a 4°C soak. The PCR fragments were separated by gel electrophoresis with 1.5% agarose gel in 0.5× TAE buffer. The sequencing of PCR products was done in the Purdue Genomics Core Facility, Purdue University, West Lafayette, IN 47907 (http://www.genomics.purdue.edu/).

SNP identification, selection, and genotyping

SNP markers were identified and selected via three approaches: (1) new SNP markers identified specifically for *3gG2* allele at *Rsv1* locus and for EST AW307114 linked to *Rsv4*, (2) gene-linked SNP markers selected from clone N11PF linked to *Rsv1* and from probe A519 linked to *Rsv3*, and (3) SNP markers screened from the soybean chromosome 13 (LG F) near *Rsv1*, chromosome 14 (LG B2) near *Rsv3*, and chromosome 2 (LG D1b) near *Rsv4*.

A *3gG2* allele-specific primer pair 3gG2-f1 (TCAGC AGAATCATAGCCACG)/3gG2-r1 (CCAACACCACTTG CATTTTG) was designed from the sequence of the GenBank accession AY518517 using the primer design tool—Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). AY518517 is an accession in GenBank that contains the sequence of the *Glycine max* disease resistance protein (*3gG2*) gene, which is the active allele at the *Rsv1* locus for SMV resistance (Hayes et al. 2004). The forward and reversal primers are located at 2257–2276 base and 2940–2921 base of AY518517, respectively. The *3gG2* allele-derived SNP markers were identified from the sequences amplified from the primer pair 3gG2-f1/r1 between the susceptible cultivar Essex and soybean genotypes that contained the *3gG2* allele at *Rsv1* locus by multiplex sequence alignment using the software BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).

AW307114 allele-specific primer pair AW307114-f1 (AATGCCCACAATCAAGGAAG)/AW307114-r1 (GAA TTTAAGGTTCCGGTTGAGA) was designed from an EST sequence (GenBank accession AW307114). An EST marker, AW307114A, developed from this EST sequence was mapped at 3.3 cM away from *Rsv4* (Hwang et al. 2006). The forward and reversal primers are located at 78–97 bases and 503–482 bases of the AW307114 sequence, respectively.

N11PF was selected for identification of *Rsv1*-linked SNP markers because it was identified to be tightly linked to *Rsv1* with a distance less than 2 cM (Gore et al. 2002; Li et al. 2004). Five SNP and one InDel markers were

Table 3 Seven SNP markers detected from the A519 allele among 2–13 sequences from GenBank

GenBank accession	Soybean genotype	A519-snp1	A519-snp2	A519-snp3	A519-snp4	A519-snp5	A519-snp6	A519-snp7
AF125925	Jackson	275C[a]	331C	433A	–	–	–	–
AF125926	Lincoln	275C	331C	433A	–	–	–	–
AF125927	Ogden	274C	330C	432A	–	–	–	–
AF125928	S100	274C	330C	432A	–	–	–	–
AF125929	Tokyo	291C	347C	449A	–	–	–	–
AF125931	Mandarin Illinois	274A	330G	432G	–	–	–	–
AF125932	PI 437654	274A	330G	432G	–	–	–	–
AF125934	A81-356022	343C	399G	501G	–	–	–	–
AF125924	A. K. Harrow	274C	330C	432A	452C	534T	–	–
AF125930	CNS	274A	330G	432G	481T	563C	–	–
AF125933	BSR-101	303C	359G	461G	481T	563T	–	–
AF348331	L29	302A	358G	460G	480T	562C	956C	1165G
AF348332	Lee68	304C	360C	462A	482C	564T	958G	1163A

[a] The SNP location and type, e.g. 275C signifies that the SNP A519-snp1 has a 'C' base at location of 275 site of the AF125925 sequence

Table 4 Twenty-two SNP markers and their positions in relation to three SMV resistance loci on soybean molecular linkage group (MLG)

SNP name/locus	SNP type	MLG	Position (cM)	Linked Gene	References
3Gg2-snp1	G/T	F	69.9	*Rsv1*	Shi et al. (2008a, b)
3Gg2-snp2	A/C	F	69.9	*Rsv1*	
N11PF-snp2	G/T	F	68.5	*Rsv1*	Gore et al. (2002)
N11PF-snp3	G/T	F	68.5	*Rsv1*	Jeong and Saghai Maroof (2004)
A519-snp2	C/G	B2	91.9	*Rsv3*	Jeong and Saghai Maroof (2004)
A519-snp4	C/T	B2	91.9	*Rsv3*	Choi et al. (2007)
AW307114-indel1	A/-	D1b	52.1	*Rsv4*	Hwang et al. (2006)
BARC-010137-00526[a]	A/G	F	68.1	*Rsv1*	Choi et al. (2007)
BARC-025897-05144	C/T	F	68.9	*Rsv1*	
BARC-022043-04271	C/T	F	69.9	*Rsv1*	
BARC-029823-06424	C/G	F	69.9	*Rsv1*	
BARC-015435-01966	A/C	F	75.8	*Rsv1*	
BARC-012953-00413	C/T	B2	90.8	*Rsv3*	
BARC-039667-07534	A/G	B2	90.8	*Rsv3*	
BARC-024363-04860	C/G	B2	94.7	*Rsv3*	
BARC-010469-00648	A/C	B2	96.8	*Rsv3*	
BARC-012703-00380	G/T	B2	96.8	*Rsv3*	
BARC-040713-07825	A/G	D1b	42.7	*Rsv4*	
BARC-025955-05182	A/G	D1b	46	*Rsv4*	
BARC-025183-06458	A/T	D1b	51.9	*Rsv4*	
BARC-011149-00860	A/d	D1b	68.8	*Rsv4*	
BARC-011147-00855	A/G	D1b	71.53	*Rsv4*	

[a] All BARC series SNP markers were from http://bfgl.anri.barc.usda.gov/soybean/index.html (Choi et al. (2007))

postulated from eight N11PF fragments by BioEdit and are listed in Table 2. N11PF-snp2 and N11PF-snp3 were confirmed to be polymorphic by AS-PCR in parents and two F_2 populations derived from PI96983 × Lee68 and 'Myeongjunamulkong' × Lee68, respectively (Jeong and Saghai Maroof 2004) and thus they were selected for this research. A519 was selected for identification of *Rsv3*-linked SNP markers. A519 is a RFLP probe which was reported to be linked to *Rsv3* with a distance of 0.8 cM (Jeong et al. 2002) and 13 sequences on the A519 allele

Table 5 Primer properties in a 11-SNP iPlex assay for Sequenom SNP genotyping

SNP_ID	2nd-PCRP	1st-PCRP	UEP_mass	UEP_SEQ	Linked gene
3gG2-snp1	ACGTTGGATGTTTGATCAGTTCATGAACCC	ACGTTGGATGCATGGTAGGAAAGGGAGTTG	6654.3	ATCAGTTCATGAACCCTGCACT	*Rsv1*
3gG2-snp2	ACGTTGGATGAAGGACCTGTTCCATCCTTG	ACGTTGGATGCTGCACTTGCAAGGTACCAA	5540.6	ACCAACAGGAGAAGATTT	*Rsv1*
N11PF-snp2	ACGTTGGATGAGAGCATGACCAATCCAAAG	ACGTTGGATGCCCCCTCTTTTCTTCCTAATC	7730.1	AGAAACTTTAAAACTGGGTAATAAA	*Rsv1*
N11PF-snp3	ACGTTGGATGTTGGATGGTTAAAGATGTG	ACGTTGGATGCAACTTGTACAGATAAAGCC	6524.3	ATGGTTAAAGATGTGTGACTA	*Rsv1*
Barc-015435-01966	ACGTTGGATGAGATGTCATAAATTACTCC	ACGTTGGATGGGAAGGCAGATCTGGAATTG	5059.3	AACACTTTCGCCCCCAA	*Rsv1*
Barc-012953-00413	ACGTTGGATGAACTTTGAGCTCTTGCCTCC	ACGTTGGATGATGTGCAAGCTAAACTGCCC	5139.4	CAGTCCCAGATTGACAC	*Rsv3*
A519-snp2	ACGTTGGATGTGGCCTCTGAGGCAATTAAG	ACGTTGGATGTAGTACCTCTCACAATAGC	5931.9	TGAGGCAATTAAGGTTGAG	*Rsv3*
A519-snp4	ACGTTGGATGTTCTTTGTTCACAAACCAGG	ACGTTGGATGGAAACATGTCCAACTATTACC	7622	TCTTTGTTCACAAACCAGGTTAGTT	*Rsv3*
Barc-025955-05182	ACGTTGGATGGAACTCATCAATGCCTGCAC	ACGTTGGATGTTGGTACCATGCCAAGGAG	6077	ATGCCTGCACTGATCCATGA	*Rsv4*
Barc-011147-00855	ACGTTGGATGTTACATCAGTTGTTCCTTC	ACGTTGGATGAAACGACCTAGGATCACTTC	6442.2	AGTTGTTCCTTCATAGGTGAT	*Rsv4*
AW307114snp1	ACGTTGGATGTTCCGATTCTCATCTGCGTC	ACGTTGGATGCTATAAATTTATTCTTCCGC	7021.6	TTTGTCTTTACCTGGCAAAAAAA	*Rsv4*

were published in the GenBank. Seven SNPs were found from the 13 sequences by alignment using BioEdit (Table 3). The A519-snp3, A519-snp6, and A519-snp7 in Table 3 have been verified by Jeong and Saghai Maroof (2004) by comparing the sequences of A519 alleles between L29 and Lee 68 and by AS-PCR. The A519-snp2 and A519-snp4 were selected for detecting *Rsv3* in this research. In addition, four SNP loci from chromosome 13 (LG F) near the *Rsv1* locus, five from MLF B2 near *Rsv3*, and four from chromosome 2 (LG D1b) near *Rsv4* based on the genetic map in soybean (Choi et al. 2007; http://bfgl.anri.barc.usda.gov/soybean/index.html) were selected for this research (Table 4). An AW307114 allele-derived SNP (InDel) marker was identified and also selected as one of the *Rsv4*-linked markers in this research. After screening SNP markers in diverse soybean germplasm, an eleven SNP multiplex assay consisting of two *3gG2*-derived, two A519 allele-derived, one AW307114 allele-derived, and six selected from chromosome 13 (LG F), 14 (B2), and 2 (D1b) was designed for Sequenom iPLEX genotyping; the primers and the multiplex assay were listed in Table 5 consisting of 10 SNPs plus one InDel iPLEX, which was designed by the tool—Assay Design 3 specific for iPLEX genotyping on MassARRAY system by Sequenom; the software Assay Design can be downloaded from Sequnom web site (https://www.mysequenom.com/) after registration.

SNP genotyping was done at the Sequenom Technology Core, the Division of Human Genetics in the Department of Genetics, Washington University, St. Louis, MO 63110 (http://hg.wustl.edu/info/Sequenom_description.html) on the MassARRAY system using iPLEX GOLD. The processing of iPLEX GOLD reactions included iPLEX PCR, SAP, and iPLEX Extend following the iPLEX Gold Application Guide from the Sequenom (http://www.sequenom.com/Files/Genetic-Analysis—Graphics/iPLEX-Application-PDFs/iPLEX-Gold-Application-Guide-v2r1) (Gabriel et al. 2009; Masouleh et al. 2009; Pattemore and Henry 2008). The procedure of iPLEX PCR is the same as the normal PCR. Briefly, 10 ng genomic DNA was amplified in a 5 μl reaction containing 1 × HotStar Taq PCR buffer (Qiagen), 1.625 mM MgCl2, 0.5 mM each dNTP, 0.1 μM each PCR primer, and 0.5 U HotStar Taq DNA polymerase (Qiagen). The reaction was incubated at 94°C for 4 min followed by 45 cycles of 94°C for 20 s, 56°C for 30 s, 72°C for 1 min, and then followed by 3 min at 72°C. After iPLEX, excess dNTPs were removed from the reaction by added 2 μl shrimp alkaline phosphatase (SAP) enzyme solution (1.53 μl water (HPLC grade), 0.17 μl SAP buffer (10×), 0.30 μl SAP enzyme (1.7 U/μl)) into each sample well and mixed, and then incubated at 37°C for 20 min followed by 5 min at 85°C to deactivate the enzyme—called SAP procedure in iPLEX. Then, the

Table 6 Two SNP markers detected from the *3gG2* allele at *Rsv1* locus for SMV resistance

Genotype	GenBank accession	3gG2-snp1		3gG2-snp2	
		21 bp sequence	Base	21 bp sequence	Base
PI96983	AY518517	ACCCTGCACTTCACTTCAAAT	2415T[a]	CAAATTTCTTCAAATCTTCTC	2815C
PI96983	GQ496314	ACCCTGCACTTCACTTCAAAT	159T	CAAATTTCTTCAAATCTTCTC	559C
J05	GQ496315	ACCCTGCACTTCACTTCAAAT	159T	CAAATTTCTTCAAATCTTCTC	559C
Marshall	GQ496312	ACCCTGCACTTCACTTCAAAT	159T	CAAATTTCTTCAAATCTTCTC	559C
Ogden	GQ496313	ACCCTGCACTTCACTTCAAAT	159T	CAAATTTCTTCAAATCTTCTC	559C
York	GQ496317	ACCCTGCACTGCACTTCAAAT	159G	CAAATTTCTTAAAATCTTCTC	559C
V94-5152	GQ496318	ACCCTGCACTGCACTTCAAAT	159G	CAAATTTCTTAAAATCTTCTC	559A
Essex	GQ496316	ACCCTGCACTGCACTTCAAAT	159G	CAAATTTCTTAAAATCTTCTC	559A

[a] The SNP location and type, e.g. 3gG2-snp1 is located at 2415 base and 3gG2-snp2 at 2815 base of the GenBank accession AY518517 which is the *Glycine max* disease resistance protein (*3gG2*) gene at *Rsv1* locus for SMV resistance (Hayes et al. 2004)

iPLEX Extend was carried out with a final concentration of between 0.625 and 1.5 μM for each extension primer, depending on the mass of the probe, iPLEX termination mix (Sequenom) and 1.35 U iPLEX enzyme (Sequenom) and conducted a two-step cycles program; 94°C for 30 s followed by 40 cycles of 94°C for 5 s, then followed 5 cycles of 52°C for 5 s, and 80°C for 5 s within the 40 cycles, then 72°C for 3 min in the 40 cycles. The reaction was then desalted by addition of 6 mg resin to each well followed by mixing and centrifugation to settle the contents of the tube. The extension product was spotted onto a 96-well spectroCHIP before being flown in the MALDI-TOF mass spectrometer (Sequenom).

Data analysis

Segregation ratios for SMV reaction and the molecular data of the segregating populations derived from J05 × Essex were tested for goodness-of-fit and independence using a chi-square test (Liu 1998). Linkage analysis was performed based on the maximum likelihood estimator. The recombination fraction (r) was calculated using a SAS script (SAS Institute, Cary, NC), which was kindly provided by Dr. Ben-Hui Liu, Department of Forestry at North Carolina State University (Raleigh, NC) and Plexigen Inc. (Shi and Qiu 2000). The recombination fraction between loci was transformed according to the Kosambi function using the formula $C_{AB} = 1/4\ln\ [(1 + 2r)/(1 - 2r)] \times 100$ (Weir 1996), where C_{AB} is the map distance (cM), and r is the estimated recombination fraction.

Results

Primer pair 3gG2-f1/r1 produced a 684 bp DNA fragment in seven soybean lines, i.e. PI 96983, Marshall, Ogden, J05, York, V92-5152, and Essex. The corresponding sequences have been submitted to GenBank with the accessions number from GQ496312 to GQ496318. Two SNPs, 3gG2-snp1 and 3gG2-snp2 were observed from these sequences by multiple sequence alignment. The SNP 3gG2-snp1 [G/T] was located at 2415 base of AY518517 and 159 base of GQ496312 to GQ496318, and the SNP 3gG2-snp2 (C/A) located at 2815 base of AY518517 and 559 base of GQ496312 to GQ496318 (Table 6).

The primer pair AW307114-f1/r1 produced a DNA fragment with the size of 690 or 691 bp from J05, V94-5152, PI 486335, L29, Jindou1, and Essex. The sequences were published in GenBank with the accessions numbers from GQ496306 to GQ496311. An InDel [A/d] marker was found in the PCR fragments amplified from AW307114f1/r1. The base [A] was observed at 194 base of the sequences amplified from J05, V94-5152, and PI 486335 (GenBank accessions GQ496306, GQ496307, and GQ496308), but the deletion [d] was found between 193 and 194 base of the sequences amplified from Essex, L29, and Jindou1 (GenBank accessions GQ496311, GQ496310, and GQ496309). In comparison with the EST AW307114 sequence, all PCR fragments amplified from AW307114-f1/r1 had an intron with 265 bp length from 73 to 337 bases of the PCR fragments and the corresponding site in AW307114 was between 149 and 150 bases of the sequence. After removing the intron, the multiple sequence alignment among the PCR fragments and the corresponding sequences of AW307114 showed that all sequences were the same except for the base at 450 base which is A in the AW307114, but C in the other six sequences. There appeared to be a SNP at this site, but this needs to be further validated.

Eleven SNP markers showed polymorphic among 47 diverse soybean genotypes in this research (Table 1). Most of soybean showed homogenous SNP type such as [T T] or [G G], but heterogeneous SNP such as [G T] was also observed. However, the heterogeneous SNP type [G T] did

Table 7 Co-segregation of the SNP markers and SMV resistance allele *Rsv1* in the F_2 population derived from J05 × Essex

Locus		Phenotype[a]						χ^2_A (3:1)	χ^2_B (1:2:1)	χ^2_{AB} (3:6:3:1:2:1)
Locus A	Locus B	RM11	RM12	RM22	SM11	SM12	SM22			
Rsv1	3gG2snp2	19	55	0	0	0	20	0.181	2.836	265.939**
Rsv1	A519-snp2	18	40	17	2	12	5	0.181	1.149	3.362 ND
Rsv1	A519-snp4	18	40	17	2	12	5	0.181	1.149	3.362 ND
Rsv1	AW307114-indel1	24	35	16	5	11	3	0.181	2.17	4.042 ND

χ^2_A: Chi-square test for the phenotypic segregations on G1 inoculations

χ^2_B: Chi-square test for the genotypic segregation of SNP marker

χ^2_{AB}: Chi-square test for the combination of phenotypic segregations on G1 inoculations and genotypic segregation of SNP marker

** Significantly different from expected ratio at $P = 0.01$, and 'ND' not significantly different from expected ratio at $P = 0.01$

[a] Phenotype: RM11, SMV resistant and homozygous SNP allele one; RM12, SMV resistant and heterozygous SNP type; RM22, SMV resistant and homozygous SNP allele two; SM11, SMV susceptible and homozygous SNP allele one; SM12, SMV susceptible and heterozygous SNP type; and SM22, SMV susceptible and homozygous SNP allele two

Table 8 Co-segregation of the SNP markers and SMV resistance allele *Rsv3* for resistance to G7 in the SMV-G1-susceptible F_2 population derived from J05 × Essex based on phenotypes of the $F_{2:3}$ lines

Locus		Phenotype[a]						χ^2_A (3:1)	χ^2_B (1:2:1)	χ^2_{AB} (3:6:3:1:2:1)
Locus A	Locus B	RM11	RM12	RM22	SM11	SM12	SM22			
Rsv3	3gG2snp2	0	0	34	0	0	14	0.341	0.00 (0:1)	–
Rsv3	A519-snp2	11	23	0	0	0	14	0.341	0.337	60.166**
Rsv3	A519-snp4	11	23	0	0	0	14	0.341	0.337	60.166**
Rsv3	AW307114-indel1	13	15	6	4	8	2	0.341	3.458	0.604 ND

χ^2_A: Chi-square test for the phenotypic segregations on G7 inoculations

χ^2_B: Chi-square test for the genotypic segregation of SNP marker

χ^2_{AB}: Chi-square test for the combination of phenotypic segregations on G7 inoculations and genotypic segregation of SNP marker

** Significantly different from expected ratio at $P = 0.01$, and 'ND' not significantly different from expected ratio at $P = 0.01$

[a] Phenotype: RM11, SMV resistant and homozygous SNP allele one; RM12, SMV resistant and heterozygous SNP type; RM22, SMV resistant and homozygous SNP allele two; SM11, SMV susceptible and homozygous SNP allele one; SM12, SMV susceptible and heterozygous SNP type; and SM22, SMV susceptible and homozygous SNP allele two

not necessarily mean a heterogeneous locus for the SNP site, it can be caused by analogous or paralogous alleles which have the same sequences of the primer pair but has different base at the SNP site. Thus the heterogeneous SNP type was amplified from two alleles at the same PCR reaction. The SNP type [0 0] signifies no data obtained and may be caused by experiment error or by a mutation on the primer sequences for a particular soybean genotype. The SNP marker 3gG2-snp-1 had three alternative SNP types, [G G], [G T], or [T T], and 3gG2-snp2 also showed three alternative types, [A A], [C C], or [A C]. The haplotypes of 3gG2-snp1 and 3gG2-snp2 mainly consisted of four types: [T T], [C C]; [G G], [C C]; [G T], [A A] and [G T], [C C]. All soybean genotypes that contained the *Rsv1* alleles except for those with *Rsv1-y* or with [0 0] data expressed a haplotype of [T T], [C C]. All soybean genotypes that contained the *Rsv1-y* i.e. 'Davis', 'York', and 'Zho 18' expressed a haplotype of [G G], [C C]. And those without *Rsv1* or *Rsv1* alleles, mainly showed two haplotypes: 'Da bai ma' carrying *Rsv3+4* (Zheng et al. 2008), 'Essex' with no SMV resistance allele (rsv) (Chen et al. 1991; Zheng et al. 2008), 'V229' carrying *Rsv3*, 'V94-5152', 'V-94-9003', and 'Virginia' carrying *Rsv4* (Zheng et al. 2005, 2008) showed haplotype of [G T], [A A]; and 'Columbia' carrying *Rsv3+4*, L29 carrying *Rsv3* had haplotype of [G T], [C C]. This result was similar to that of a previous study where a *3gG2* allele-derived PCR-based marker Rsv1-f/r was used as a marker specific for the SMV resistance allele at the *Rsv1* locus (Shi et al. 2008b). Besides these for haplotypes, the cultivar 'Kaeri-gnt 220-7' showed a different one from the others indicating that genetic diversity and mutation may have existed in the 3gG2 gene supporting the observation that the *Rsv1* locus was a multigene family in which closely linked members contribute to the host-pathogen interaction (Hayes et al. 2004). According to these results, an adventurous deduction was

made that 'Ching Tao 21' and 'Suzuyataka' will carry one of the *Rsv1* alleles except for *Rsv1-y* because they have the haplotype of [T T], [C C].

In our validation study, the 'J05' × 'Essex' F_2 segregation for SMV-G1 resistance (*Rsv1*) phenotype fits 3 resistant: 1 susceptible as expected for a single dominant gene; and the F_2 segregation for the four SNP makers, 3gG2snp2, A519-snp2, A519-snp4, and AW307114-indel1, each fits a 1:2:1 ratio as expected for a single co-dominant marker (Table 7). The SMV resistance to G1 in J05 was previously shown to be controlled by the dominant gene *Rsv1* and confirmed by SSR markers (Zheng et al. 2006; Shi et al. 2008a). The four SNP markers all showed co-dominant alleles for their corresponding locus. Therefore, there are six possible combinations between phenotypic (R, resistant; S, susceptible) and the co-dominant marker (M11, M12, M22) in the F_2 population: RM11, SMV resistant and homozygous SNP type for allele one; RM12, SMV resistant and heterozygous SNP type; RM22, SMV resistant and homozygous SNP type for allele two; SM11, SMV susceptible and homozygous SNP type for allele one; SM12, SMV susceptible and heterozygous SNP type; and SM22, SMV susceptible and homozygous SNP type for allele two. The co-segregation ratio of RM11:RM12:RM22:SM11:SM12:SM22 between the SNP marker 3gG2-snp2 and the resistance allele (*Rsv1*) did not fit 3:6:3:1:2:1 for an independent segregation of a dominant gene and a co-dominant SNP marker (Table 7), indicating 3gG2-snp2 was linked to the resistance allele (*Rsv1*) and the distance between them was calculated to be 0.0 cM (Table 7), confirming the result reported by Hayes et al. (2004). The co-segregation ratios of RM11:RM12:RM22:SM11:SM12:SM22 between other SNP markers, A519-snp2, A519-snp4, and AW307114-indel1 and the resistance allele (*Rsv1*) fit 3:6:3:1:2:1 for an independent segregation for the dominant gene and a co-dominant marker (Table 7), indicating that three SNP markers were located at a different MLG apart from the resistance gene *Rsv1*. Because the probe A519 was reported to be linked to *Rsv3* on chromosome 14 (LG B2) (Jeong et al. 2002) and EST AW307114 was reported to be linked to *Rsv4* on Chromosome 2 (LG D1b) (Hwang et al. 2006), the resistance gene for G1 in 'J05' was not located at chromosome 14 (LG B2), nor Chromosome 2 (LG D1b). These results further confirmed that the gene for resistance to G1 in 'J05' was *Rsv1*, not *Rsv3*, nor *Rsv4*. These results also demonstrated that these SNP markers can be used to identify SMV resistance genes in a segregating population.

Rsv1 confers resistance to G1 and a necrotic reaction to G7, whereas *Rsv3* confers resistance to G7 but a susceptible reaction to G1. Therefore, *Rsv1* in 'J05' can be identified in the 'Essex' × 'J05' F_2 population when inoculated with G1. Then, the G1-susceptible plants in the 'Essex' × 'J05' F_2 population should contain only *Rsv3*, but not *Rsv1* and thus can be used as a segregating population for *Rsv3*. The individual G1-susceptible plants in the 'Essex' × 'J05' F_2 population were harvested to develop $F_{2:3}$ lines. The derived $F_{2:3}$ lines were inoculated with G7 and the results showed a 1:2:1 ratio for a dominant gene (Shi et al. 2008a). In this study, the SNP genotyping was conducted in the G1-susceptible plants from the 'Essex' × 'J05' F_2 population for confirmation of *Rsv3* in 'J05' and the phenotypic data for SMV reaction to G7 in $F_{2:3}$ lines were used to figure out the genotypes of the G1-susceptible F_2 plants for the *Rsv3* locus. Results showed that the segregation for SMV resistance (*Rsv3*) phenotype fits a 3:1 ratio as expected for a single dominant gene and that the F_2 segregation for the three SNP makers, A519-snp2, A519-snp4, and AW307114-indel1 each fit a 1:2:1 ratio as expected for a single co-dominant marker (Table 8). Because all F_2 plants were G1-susceptible, the SNP marker 3gG2-snp2 for the *Rsv1* gene showed the same SNP type [A A] as 'Essex' (Table 8). The co-segregation ratio of RM11:RM12:RM22:SM11:SM12:SM22 between the SNP markers, A519-snp2, A519-snp4 and the resistance allele (*Rsv3*) did not fit the ratio of 3:6:3:1:2:1 for an independent segregation of a dominant gene and a co-dominant marker (Table 8). Further analysis showed that the two SNP markers were linked to resistance allele (*Rsv3*) and the distance was calculated to be less than 0.1 cM. The co-segregation ratio of RM11:RM12:RM22:SM11:SM12:SM22 between the SNP marker AW307114-indel1 and the resistance allele (*Rsv3*) fit 3:6:3:1:2:1 for an independent segregation of a dominant gene and a co-dominant marker (Table 8), indicating that the resistance gene was located on different MLG apart from the AW307114-indel1. A519 was reported to be linked to *Rsv3* with a distance 0.9 cM based on the data from 195 $F_{2:3}$ lines of 'L29' × 'Lee68' (Jeong et al. 2002) and the EST AW307114 was reported to be linked to *Rsv4* with a distance 3.3 cM on Chromosome 2 (LG D1b) (Hwang et al. 2006). Therefore, the gene for resistance to G7 in 'J05' was further confirmed to be *Rsv3*, not *Rsv1*, nor *Rsv4* by SNP markers in this study.

Discussion

Eleven SNP/InDel markers were tested for three genes of SMV resistance in this research. The SNP markers, 3gG2-snp1, 3gG2-snp2, N11PF-snp2, N11PF-snp3, and Barc-015435-01966 are *Rsv1*-linked; A519-snp2, A519-snp4, and Barc-012953-00413 are *Rsv3*-linked; and Barc-011147-00855, Barc-025955-05182, and AW307114-snp1 are *Rsv4*-linked (Hayes et al. 2004; Jeong and Saghai Maroof 2004; Hwang et al. 2006; Choi et al. 2007;

http://bfgl.anri.barc.usda.gov/soybean/index.html). 3gG2-snp1 and 3gG2-snp2 were developed from the candidate gene *3gG2* at *Rsv1* locus (Hayes et al. 2004) and verified to be co-segregated with the *Rsv1* gene in the Essex (rsv) × J05 (*Rsv1, Rsv3*) (Zheng et al. 2006; Shi et al. 2008a) F_2 population in this research, which further confirmed the *3gG2* was located at the *Rsv1* locus (Hayes et al. 2004). The combination of the two SNP markers 3gG2-snp1 and 3gG2-snp2 showed to be useful to identify the *Rsv1* gene in soybean genotypes: the combination of [T T] [C C] of the SNP markers were observed specific for all genotypes tested that carry *Rsv1* alleles (*Rsv1, Rsv1-h, Rsv1-k, Rsv1-m, Rsv1-n, Rsv1-r*, and *Rsv1-t*) except *Rsv1-y* (*Rsv1-s* was not been tested) (Table 1), but the combination of [G G][C C] was specific for *Rsv1-y* gene and observed in all tested soybean genotypes carrying *Rsv1-y* such as 'Davis', 'York', and 'Zho 18'. N11PF-snp2 and N11PF-snp3 have been confirmed to be polymorphic by AS-PCR (Jeong and Saghai Maroof 2004). The probe A519 was reported to be linked to *Rsv3* on chromosome 14 (LG B2) (Jeong et al. 2002). The two SNP markers, A519-snp2 and A519-snp4 were developed from the A519 were confirmed to be tightly linked the *Rsv3* in J05 based on the data from the Essex × J05 $F_{2:3}$ population in this research. EST AW307114 was reported to be tightly linked to *Rsv4* on Chromosome 2 (LG D1b) (Hwang et al. 2006). The Insert and deletion marker, AW307114-indel1 was developed from AW307114 and confirmed it was independent to Rsv1 and Rsv3 based on the Essex × J05 populations in this study. Therefore, these markers can be potentially useful in identification of specific SMV resistance genes in soybean germplasm and in selection of individuals carrying specific gene combinations in a breeding population. A multiplex assay consisting of these 11 SNP/InDel markers can be used to investigate the mode of inheritance in a SMV resistant soybean line carrying *Rsv1*, *Rsv3*, or *Rsv4* through a segregating population with phenotypic data, and to select a specific gene or pyramid two or three genes for SMV resistance through MAS in soybean breeding program. For example, the SMV resistance gene(s) in 'Jiunong 21' soybean can be identified using this 11 SNP/InDel assay in a segregating population such as F_2 from 'Essex' × 'Jiunong 21'. 'Jiunong 21' was reported to be resistant to G1–G7 except G4 (G4 was not available) and postulated to carry *Rsv1-h, Rsv1Rsv3*, or *Rsv1Rsv4* (Zheng et al. 2008). The SNP markers 3gG2-snp2 and Barc-015435-01966 can be used to examine whether 'Jiunong 21' carries *Rsv1* because 'Essex' would have [A A] but 'Jiunong 21' would have [C C] in 3gG2-snp2, and Essex would have [C C] but 'Jiunong 21' would have [A A] in Barc-015435-01966. Likewise, the *Rsv3*-linked SNP markers, A519-snp2 and A519-snp4 and the *Rsv4*-linked SNP marker Barc-025955-05182 can used to test whether 'Jiunong 21' carries *Rsv3* and/or *Rsv4* because the three SNP markers were polymorphic between 'Essex' and 'Jiunong 21' (Table 1). As an example for gene pyramiding, the *Rsv4* in 'V94-9003' can be combined with *Rsv1-y* and *Rsv3* in 'Zho18' because the *Rsv1*-linked SNP markers, 3gG2-snp1 and 3gG2-snp2, the *Rsv3*-linked SNP markers, A519-snp2 and A519-snp4, and the *Rsv4*-linked SNP markers, Barc-011147-00855 and AW307114-indel1 showed polymorphism between 'V94-9003' (*Rsv4*) and 'Zho 18' (*Rsv1-yRsv3*) (Table 1). Research is in progress to validate these SNP markers and their utility in a MAS scheme. At this time, the 11 SNP/InDel multiplex assays can only differentiate *Rsv1-y* from other alleles at the *Rsv1* locus. With the allele-specific SNP discovery, it is expected that a new multiplex can be developed to differentiate alleles, *Rsv1, Rsv1-h, Rsv1-k, Rsv1-m, Rsv1-n, Rsv1-r, Rsv1-s, Rsv1-t, Rsv1-y* at *Rsv1* locus.

In this study, we provide an effective approach to identify *3gG2* gene-derived SNP markers and EST AW307114 gene-linked SNP marker for SMV resistance alleles at *Rsv1* and *Rsv4* locus. We also identified gene-linked SNP markers for SMV resistance genes, such as N11PF allele-derived SNP markers linked to *Rsv1* and A519 allele-derived SNP markers linked to *Rsv3*. A genome region specific screen was also conducted to identify linked SNP markers for *Rsv1*, *Rsv3*, and *Rsv4* from soybean composite genetic maps. An 11-SNP/InDel multiplex assay was developed for selection of SMV resistance genes *Rsv1*, *Rsv3*, and *Rsv4* and it will provide breeders with a tool for selecting these genes through MAS in a soybean breeding program.

References

Brahm L, Rocher T, Friedt W (2000) PCR-based markers facilitating marker assisted selection in sunflower for resistance to downy mildew. Crop Sci 40:676–682

Caicedo AL, Williamson SH, Hernandez RD (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS genetics 3:1745–1756

Chen P, Choi CW (2008) Characterization of genetic interaction between soybean and soybean mosaic virus. In: Rao GP (ed) Molecular diagnosis of plant viruses. Studium Press, Houston TX, pp 389–422

Chen P, Buss GR, Roane CW, Tolin SA (1991) Allelism among genes for resistance to soybean mosaic virus in strain differential soybean cultivars. Crop Sci 31:305–309

Chen P, Ma G, Buss GR, Gunduz I, Roane CW, Tolin SA (2001) Inheritance and allelism tests of Raiden soybean for resistance to soybean mosaic virus. J Hered 92:51–55

Cho EK, Goodman RM (1979) Strains of soybean mosaic virus: classification based on virulence in resistant soybean cultivars. Phytopathology 69:467–470

Choi I, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, Hwang EY, Yi SI,

Young ND, Shoemaker RC, van Tassell CP, Specht JE, Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and single nucleotide polymorphism analysis. Genetics 176:685–696

Fjellstrom R, Conaway-Bormans CA, McClung AM, Marchetti MA, Shank AR, Park WD (2004) Development of DNA markers suitable for marker assisted selection of three *Pi* genes conferring resistance to multiple *Pyricularia grisea* pathotypes. Crop Sci 44:1790–1798

Foolad MR, Sharma A (2005) Molecular markers as selection tools in tomato breeding. Acta Hort 695:225–240

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the sequenom massARRAY iPLEX platform. Current protocols in human genetics, vol 60, pp 2.12.1–2.12.18. John Wiley & Sons, Inc, NY

Giancola S, McKhann HI, Berard A, Camilleri C, Durand S, Libeau P, Roux F, Rebound X, Gut IG, Brunel D (2006) Utilization of three high-throughput SNP genotyping methods, the GOOD assay, Amplifluor and Taqman, in diploid and polyploidy plants. Theor Appl Genet 112:1115–1124

Gore MA, Hayes AJ, Jeong SC, Yu 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

Hayes AJ, Saghai Maroof MA (2000) Targeted resistance gene mapping in soybean using modified AFLPs. Theor Appl Genet 100:1279–1283

Hayes AJ, Ma G, Buss GR, Saghai Maroof MA (2000) Molecular marker mapping of *Rsv4*, a gene conferring resistance to all known strains of soybean mosaic virus. Crop Sci 40:1434–1437

Hayes AJ, Jeong SC, Gore MA, Yu YG, Buss GR, Tolin SA, Saghai Maroof MA (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

Hwang TY, Yu S, Yang K, Kim HM, Jeong SC (2006) Application of comparative genomics in developing molecular markers tightly linked to the virus resistance gene *Rsv4* in soybean. Genome 49:380–388

Hyten DL, Hartman GL, Nelson RL, Frederick RD, Concibido VC, Narvel JM, Cregan PB (2007) Map location of the *Rpp1* locus that confers resistance to soybean rust in soybean. Crop Sci 47:837–838

Jeong SC, Saghai Maroof MA (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, Hayes AJ, Biyashev RM, Saghai Maroof MA (2001) Diversity and evolution of a non-TIR-NBS sequence family that clusters to a chromosomal "hotspot" for disease resistant genes in soybean. Theor Appl Genet 103:406–414

Jeong SC, Kristipati S, Hayes AJ, Maughan PJ, Noffsinger SL, Gunduz I, Buss RG, 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

Kamiya M, Kiguchi T (2003) Rapid DNA extraction method from soybean seeds. Breed Sci 53:277–279

Kisha T, Sneller CH, Diers BW (1997) Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci 37:1317–1325

Li Z, Bernard RL, Nelson RL (1998) A RAPD marker with a heterozygous form linked to the Rsv1 locus in soybean. In: Agronomy abstracts 1998 annual meetings. USA, Baltimore, Maryland, 157

Li Z, Nelson RL, Bernard RL, Peng Y (2004) Identification and sequence analysis of a RAPD marker tightly linked to the *Rsv1* locus in soybean. Soybean Newsl 31

Liu B (1998) Statistical genomics: linkage, mapping, and QTL analysis. CRC Press, Boca Raton, FL

Masouleh AK, Waters DL, Reinke RF, Henry RJ (2009) A high-throughput assay for rapid and simultaneous analysis of perfect markers for important quality and agronomic traits in rice using multiplexed MALDI-TOF mass spectrometry. Plant Biotechnol J 7:355–363

Missaoui AM, Ha BK, Phillips DV, Boerma HR (2007) Single nucleotide polymorphism detection of the *Rcs3* gene for resistance to frogeye leaf spot in soybean. Crop Sci 47:1681–1690

Pattemore J, Henry RJ (2008) Sequenom MassARRAY iPLEX Gold SNP genotyping for high throughput variety identification. Inter plant and animal genome XVI conference, sequenome workshop, San Diego CA, USA, January 12–16

Rafalski JA (2002) Application of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Saghai Maroof MA, 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*. The Plant Genome 3:14–22

Shi A, Qiu W (2000) Plant resistance to fungus and bacterial diseases and molecular biology technology. In: Yang Z, Hu S, Zhou X (eds) Biotechnology and sustainable agriculture shanghai scientific and technological. Literature Publishing House, Shanghai, China, pp 70–100

Shi A, Chen P, Li D, Zheng C, Hou A, Zhang B (2008a) Genetic confirmation of two independent genes for resistance to soybean mosaic virus in J05 soybean using SSR markers. J Hered 99:598–603

Shi A, Chen P, Zheng C, Hou A, Zhang B (2008b) A PCR-based marker for the *Rsv1* locus conferring resistance to soybean mosaic virus. Crop Sci 48:262–268

Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer Assoc, Inc, Sunderland, MA

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology 84:60–64

Zheng C, Chang R, Qiu L, Chen P, Wu X, Chen S (2003) Identification and characterization of a RAPD/SCAR marker linked to a resistance gene for soybean mosaic virus in soybean. Euphytica 132:199–210

Zheng C, Chen P, Gergerich R (2005) Characterization of resistance to soybean mosaic virus in diverse soybean germplasm. Crop Sci 45:2503–2509

Zheng C, Chen P, Gergerich R (2006) Genetic analysis of resistance to soybean mosaic virus in J05 soybean. J Hered 97:429–437

Zheng C, Chen P, Li D, Gergerich R (2008) Identification of new alleles for resistance to soybean mosaic virus in soybean. Can J Plant Path 595–603