



Fine mapping and molecular marker development of the *Sm* gene conferring resistance to gray leaf spot (*Stemphylium* spp.) in tomato

Xiaomei Su^{1,2} · Guangtao Zhu³ · Zejun Huang¹ · Xiaoxuan Wang¹ · Yanmei Guo¹ · Baoju Li¹ · Yongchen Du¹ · Wencai Yang² · Jianchang Gao¹

Received: 15 May 2018 / Accepted: 16 November 2018 / Published online: 26 November 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Key message The tomato gray leaf spot resistance gene *Sm* was fine-mapped in a 185-kb region through a map-based cloning strategy and genome-wide association study; a candidate gene was proved to be involved in *Sm*-mediated resistance through transient gene silencing.

Abstract Gray leaf spot, caused by *Stemphylium* spp., is a warm weather foliar disease in tomato (*Solanum lycopersicum* L.). Resistance against gray leaf spot is conferred by a single incompletely dominant gene (*Sm*) located on chromosome 11. This study aimed to map and identify molecular marker tightly linked to the *Sm* gene for the use of marker-assisted selection in breeding. Using an F₂ population derived from a cross between the resistant line ‘9706’ and the susceptible line ‘Heinz 1706’, the *Sm* gene was mapped to a 185-kb interval between two markers, InDel343 and InDel-FT-32 on chromosome 11, which was consistent with the result of a genome-wide association study using 289 diverse accessions. An ORF predicted in this region was proved to be involved in *Sm*-mediated resistance through transient gene silencing and seems to be a good candidate of the *Sm* locus. To clone the *Sm* gene, a bacterial artificial chromosome (BAC) library was screened and one BAC clone B80B15 containing the predicted ORF was identified. The analysis of sequence and structure characteristics demonstrated that the candidate gene was not a typical type resistance gene. Additionally, a co-dominant marker Sm-InDel, which produced a 122-bp or 140-bp fragment for resistant or susceptible alleles, respectively, was developed. This marker was validated in 289 germplasm and could be used in marker-assisted selection for gray leaf spot resistance.

Communicated by Richard G.F. Visser.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00122-018-3242-z>) contains supplementary material, which is available to authorized users.

✉ Wencai Yang
yangwencai@cau.edu.cn

✉ Jianchang Gao
gaojianchang@caas.cn

- ¹ Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of the Ministry of Agriculture, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, No. 12, Zhongguancun Nandajie, Haidian District, Beijing 100081, China
- ² Department of Vegetable Science, College of Horticulture, China Agricultural University, No. 2 Yuanmingyuan Xilu, Beijing 100193, China
- ³ The CAAS-YNU Joint Academy of Potato Sciences, Yunnan Normal University, Kunming 650500, China

Introduction

Gray leaf spot is a warm weather foliar disease. The causal agent *Stemphylium* spp. has a wide range of hosts including important crops such as tomato (Li et al. 2010; Nasehi et al. 2012a), cucumber (Vakalounakis and Markakis 2013), garlic (Zheng et al. 2008), sweet potato (Chai et al. 2015), pepper (Kim et al. 2004; Nasehi et al. 2012b), eggplant (Nasehi et al. 2013) and spinach (Everts and Armentrout 2001; Koike et al. 2001). This disease has been widely reported in tomato in many countries including America (Hendrix and Frazier 1949), Israel (Rotem et al. 1966), Korea (Min et al. 1995; Kim et al. 1999), Venezuela (Cedeño and Carrero 1997), Malaysia (Nasehi et al. 2012a, b) and China (Li et al. 2010). Tomato gray leaf spot usually occurs and develops under warm air temperature and high humidity (Kim et al. 1999). It mainly affects the leaves of tomatoes, starting with the oldest leaves. In the early stage, small and dark spots can be seen on both surfaces of the leaves, and then, it enlarges and turns a grayish brown. Eventually, the centers of the spots

crack and fall out, and surrounding leaf areas will turn yellow and the leaves will dry and drop. Tomato gray leaf spot does not directly affect fruit, but damage to foliage reduces fruit quality and yield.

The pathogen causing tomato gray leaf spot mainly include four species: *S. solani* Weber, *S. floridanum* Hannon and Weber, *S. botryosum* Wallr., and *S. vesicarium* (Wallr.) Simmons (Bashi et al. 1973; Blancard et al. 1986). An outbreak of this disease in China was first reported in the early 2000s. More than 40% of the 30,000 greenhouses in Yutai County, Shandong Province were infected in 2002 and increased to 90% in 2003 (Liu et al. 2004). Since then, the prevalence of this disease has caused severe yield losses throughout the country (Li et al. 2010), especially in protected cultivation facilities tomatoes in spring season. A survey was performed in main tomato growing areas and two species, *S. solani* and *S. lycopersici*, were isolated from infected tissues of tomato plants growing in different seasons and localities in China. *S. solani* was identified as the epidemic species (Li et al. 2010). The fungal spores are spread by wind and water, and can survive for a long time in soil and on plant debris and weeds. Under favorable environmental conditions, *S. solani* can spread at an alarming pace, and the compatible host is devastated within several days (Li et al. 2010). It is difficult to control the disease using traditional agricultural practices and agrochemicals. Thus, the use of resistant cultivars is the most acceptable, environmentally-friendly and economically feasible method to control the disease.

Resistance to gray leaf spot was identified in *Solanum pimpinellifolium* accession PI 79532 (Andrus et al. 1942) and conferred by a single incompletely dominant gene, *Sm* (Hendrix and Frazier 1949), which is linked to the fusarium wilt race 1 resistance gene, *I*, with a crossover rate of 36% (Dennett 1950). In an incipient study, the *Sm* gene was mapped to chromosome 11 between restriction fragment length polymorphism markers TG110 and T10 at distances of 4.1 cM and 6.8 cM from the markers, respectively (Behare et al. 1991). In a recent study, Yang et al. (2017) mapped *Sm* gene in a interval of 0.26-Mb region on chromosome 11 based on bulked segregant analysis in combination with genome resequencing. However, the physical positions of the target gene from the two studies are not consistent. Until now, no further research has been reported on mapping or cloning the *Sm* gene. The most noteworthy is the *Sm* gene confers resistance to all four species of *Stemphylium*, and there are no reports of new virulent races overcoming the resistance since its introgression (Bashi et al. 1973; Blancard et al. 1986; Behare et al. 1991). Durable disease resistance is the ultimate goal of many breeding programs. There are some major R-genes, such as *Sr2* and *Lr34* in wheat, *Mlo* in barley, have proved durable over many years of agricultural use (Spielmeyer et al. 2003; Krattinger 2009; Jørgensen

1992). Therefore, cloning of *Sm* resistance gene will not only enable design of perfect markers for marker-assisted selection (MAS) in breeding, but also provide us a valuable tool to understand molecular mechanisms of host resistance against the pathogen.

Materials and methods

Plant materials

Two tomato inbred lines, ‘9706’ (KR2) and ‘Heinz 1706’, were used as parental lines to develop a segregating population for mapping of the *Sm* gene. The resistant line ‘9706’ is from the Institute of Vegetables and Flowers, and its genome has been resequenced (Lin et al. 2014). ‘Heinz 1706’ is the susceptible line for which a draft genome assembly is available (<http://solgenomics.net>). The resulting F₁ plants were self-pollinated to generate the F₂ population. A total of 441 F₂ plants were used for inheritance analysis and fine mapping of the *Sm* gene. Subsequently, a larger F₂ population, with 4730 individuals, was subjected to a recombinant screening in batches and the selected F₂ recombinants were tested for gray leaf spot resistance. To further confirm the phenotype, 315 F₃ plants derived from 15 F₂ recombinants were evaluated for gray leaf spot resistance. Each F₃ family contained 21 plants.

A collection of 289 diverse accessions previously described in detail by Lin et al. (2014) were used in a genome-wide association study (GWAS) for *Sm*. These included 145 *S. lycopersicum* accessions, 112 *S. lycopersicum* var. *cerasiforme* accessions, 28 *S. pimpinellifolium* accessions and 4 other accessions (Table S1). Three plants of each accession were tested in two independent experiments. To validate molecular marker linked to the *Sm* resistance locus, the 289 accessions were also examined using a marker-trait analysis. All of the plant materials were grown in a greenhouse in Beijing, China. The day and night average temperatures of the greenhouse were controlled at 25 °C and 20 °C, respectively.

DNA extraction and marker development

Genomic DNA was extracted from young leaves of the first 441 F₂ individuals and 289 accessions for the GWAS using the modified CTAB method (Fulton et al. 1995). For the other F₂ individuals used in the recombinant screening, DNA was extracted according to the NaOH method, as described by Wang et al. (1993) with slight modifications. In brief, small leaf samples were ground for 2 min in 120 µl of 0.4 N NaOH using a tissue striker (TissueLyser II; QIAGEN, Germany), incubated at 80 °C for 8 min, and centrifuged for 1 min at 5000×g. Then, 10 µl supernatant

was added to 200 μ l of 0.1 M Tris (pH = 8) to form (crude) genomic DNA templates for PCR.

All of the molecular markers used in this study were Insertion and Deletion (InDel) markers which were developed based on whole genome re-sequencing data of 20 cultivated tomato lines with PCR product length 80–200 bp and InDel sizes 4–6 bp. During the initial mapping stage, the InDel markers on chromosome 11 were chosen for polymorphism screening and the development of a framework map. After flanking markers were identified, more InDel markers were developed for genotyping recombinants in this region by comparing genomic DNA sequences between ‘Heinz 1706’ and ‘9706’. The genome of ‘9706’ was sequenced with $5.1 \times$ genome coverage in a previous study (Lin et al. 2014). Information on the primers for the identified markers is listed in Table S2.

Disease evaluation

The pathogenicity of two species of *Stemphylium*, 105 *S. solani* strains and 57 *S. lycopersici* strains isolated from different infected vegetables in different regions, was tested on susceptible tomato “zhongza 105” (Xie 2017). Based on the results, we choose two *S. solani* strains and two *S. lycopersici* strains in this study to test their pathogenicity on the resistant tomato line ‘9706’ and susceptible line ‘Heinz 1706’. Eventually, an isolate of *S. solani*, numbered FQ12031619, which was collected from Shandong Province, was used for inoculation. Pure isolates were grown on potato dextrose agar (PDA) plates for 7 days at 28 °C in the dark and incubated in liquid PD medium without agar in the dark at 28 °C at 160 rpm for 7 days before inoculation (Xie 2017). The whole-plant assay was performed as described by Chen (1996), with some slight modifications. In brief, plants with four fully expanded leaves were inoculated after being watered thoroughly using a sprayer to disperse the suspension (OD = 1.5–1.6) over the plants. The plants were covered with a plastic sheet for the first 2 days after inoculation to increase the humidity. Thereafter, plants were grown at 80–90% relative humidity and 24 ± 2 °C with natural light. After 7 days, the disease severity (DS) levels of plants were assessed using the severity of symptoms on a scale of 0–5: 0 = no symptoms; 1 = some older leaf area affected and showing small lesions; 2 = most leaf area affected and showing small lesions; 3 = most leaf area affected and defoliation of older leaves; 4 = most leaves wilted and/or with extensive stem damage; 5 = defoliation of the entire plant, resulting in death. For disease screening data in F_2 plants, a χ^2 -test for goodness-of-fit was used to test for deviations between observed data and the expected segregation.

Molecular marker analysis and gene mapping

Initially, seven InDel polymorphic markers well-distributed on tomato chromosome 11 based on the physical position and genetic distance were selected to construct a genetic map using 441 F_2 individuals. After the gene was located to a narrowed region, another 4730 F_2 individuals derived from the same cross were screened for recombinants using the flanking markers. Subsequently, the recombinants were analyzed with nine new markers designed in the narrowed region and their $F_{2:3}$ families were tested for gray leaf spot resistance. A linkage analysis of the *Sm* resistance locus was performed using JoinMap 4.0 (Van Ooijen 2006) with a minimum LOD threshold of 3.0.

The GWAS for *Sm*

We used the sequencing data of Lin et al. (2014) for the genome-wide association study of *Sm* gene. To eliminate sequencing and mapping errors, single nucleotide polymorphisms (SNPs) with minor allele frequencies > 5% and missing rates < 10% were used for the association analysis (Lin et al. 2014). The efficient mixed-model association expedited method was used to conduct the analysis (Kang et al. 2010).

Sequence annotation and gene expression analysis

The putative genes in the region encompassing the *Sm* locus were identified using the tomato gene model (ITAG release 3.1) in SGN (<https://solgenomics.net/>) and predicted with the online program FGENESH (<http://linux1.softberry.com/>), protein functions were predicted with BLASTP at the NCBI website (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Table S3).

Quantitative RT-PCR was performed to analyze the annotated genes expression levels in parents before and after inoculation. Total RNA was isolated from leaves of the two parental lines at 0, 2, 4 and 6 days after inoculation considering that the hyphae began infecting tomato via stomata at 2 days after inoculation (Xie 2017). Total RNAs for all of the samples were extracted using an EasyPure Plant RNA Kit (TransGen Biotech, Beijing, China). Reverse transcription was conducted using Transcript One-Step gDNA Removal and cDNA Synthesis Supermix (TransGen Biotech). The tomato actin gene (*Solyc03g078400*, F: GGGATGGAGAAG TTTGGTGGTGG and R: CTTCGACCAAGGGATGGT GTAGC) was used as an internal reference in the analyses. The qRT-PCR assay was conducted using the SYBR Green I Master Mix in a LightCycler 480 system (Roche, Switzerland). The qRT-PCRs were performed with three biological replications and three technical replicates. Gene expression levels were calculated on the basis of the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen 2001).

Virus-induced gene silencing (VIGS)

The VIGS experiment was performed with sprout vacuum-infiltration (SVI) method as described by Yan et al. (2012), an easy and effective method of agroinoculation for VIGS. Primers were designed based on predicted genes sequences to amplify 100–450 bp fragments from the cDNAs of 9706 using the Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA). PCR products were assembled into the tobacco rattle virus-based VIGS vector TRV2 and sequenced. The VIGS TRV1, TRV2 and its derivatives were transformed into *Agrobacterium tumefaciens* strain GV3101 by freeze–thaw method. A single colony was selected and inoculated in liquid culture of LB medium with antibiotics with shaking at 200 rpm at 28 °C for 12 h. The cells were re-suspended in infiltration medium (10 mmol/L MES, 10 mmol/L MgCl₂, 200 μmol/L AS, pH = 5.6) till OD₆₀₀ = 1.0 and left at room temperature for 3–4 h before agroinfiltration. *Agrobacterium* strains containing the TRV1 and TRV2 vector were mixed at a 1:1 ratio and vacuum-infiltrated into the germinating seeds of the *Sm*-resistant line 9706 using an infiltration system (a vacuum dryer connected to portable air compressor). In detail, the infiltration mixture and germinating seeds (about 21 seeds) were placed in each 10-mL beaker. And then, beakers were placed into vacuum dryer. *Agrobacterium* was infiltrated into sprouts using sprout vacuum-infiltration system set at relative vacuum degree of –25 kPa. Vacuum pressure was maintained for about 1 min and then released rapidly to atmospheric pressure. Treated sprouts were sown in pots in a growth chamber (23 ± 2 °C, 60–80% RH). The tomato phytoene desaturase gene (*PDS*) was used as the reference gene to assess the VIGS system. The TRV2 empty vector (TRV2-ev) and water were used as negative controls. Forty days after sowing, the plants were inoculated with *S. solani*. The agroinfiltration experiments were performed again with cotyledon infiltration method to confirm the candidate gene for *Sm*.

Screening of BAC library

An existing bacterial artificial chromosome (BAC), constructed with *Ty2*-resistant line 12 g-60 which was identified as resistant to *S. solani*, was screened in this study. The BAC library consisted of 99,840 clones with an average insert size of 100 kb, corresponding to 10-fold coverage of the tomato genome. The library was stored in 260 384-well microtiter plates, and all 384 clones in one plate were mixed to form a BAC pool. The BAC pool DNA was isolated by alkaline lysis method and screened by PCR using 10 primer pairs within and flanking the *Sm* region (Table S4). Afterward, the single colony from the 384-well plates corresponding to the positive pool was identified using the same markers and DNA was isolated from the positive colony.

DNA sequencing and analysis

BAC ends were sequenced to confirm that they originated from the *Sm* region, and they were also analyzed by pulsed-field gel electrophoresis to estimate the insert size. Sequence of the single-positive BAC clone harboring the predicted ORF was obtained by constructing a library of subclones (1–3 kb). Both ends of the subclones were sequenced using the ABI 3730xl platform and then assembled (BGI, Beijing, China). Putative genes in the BAC sequence were predicted with the online program FGENESH, and protein functions were predicted with BLASTP at the NCBI website.

Validation of molecular marker linked to the *Sm* gene

To verify the associations between the genotypes and phenotypes, the 289 accessions were subject to genotyping with the developed marker *Sm*-InDel (F: 5'-CTACACTTTCTCGTTCCCAATG-3', R: 5'-ATCGCCAAACCAATCAAATC-3'). The PCR was conducted with an initial denaturation step of 3 min at 94 °C, followed by 32 cycles of 30 s at 94 °C, 30 s at 55 °C and a 30 s extension at 72 °C. Final reactions were extended for 5 min at 72 °C. The PCR products were separated on 3% agarose gel, stained with ethidium bromide and photographed using a BIO-RAD Gel Documentation System (Bio-Rad Laboratories, Shanghai, China).

Results

Inheritance of gray leaf spot resistance in tomato

Previously, *Sm* has been described as an incompletely dominant gene for gray leaf spot resistance (Hendrix and Frazier 1949). To verify this conclusion, we performed the gray leaf spot disease assay on a segregating F₂ population, 12 F₁ plants and their parental lines. The resistant parent '9706' exhibited a high level of gray leaf spot resistance to *S. solani* (P₁, DS level = 0), while the other parent, 'Heinz 1706', showed lesions, wilted and died 7 days after inoculation (P₂, DS level = 5) (Fig. 1a). The F₁ plants showed small lesions on old leaves (DS level = 1 or 2). Of the 441 F₂ individuals, 340 had DS levels from 0 to 3 and were considered as resistant. The remaining 101 plants were completely or extensively blighted (having DS levels between 4 and 5) and were regarded as susceptible (Fig. 1b, c). The resistant/susceptible segregation ratio among plants was 3:1 ($\chi^2 = 1.03$, $P = 0.31$), suggesting the presence of a single dominant resistance gene.

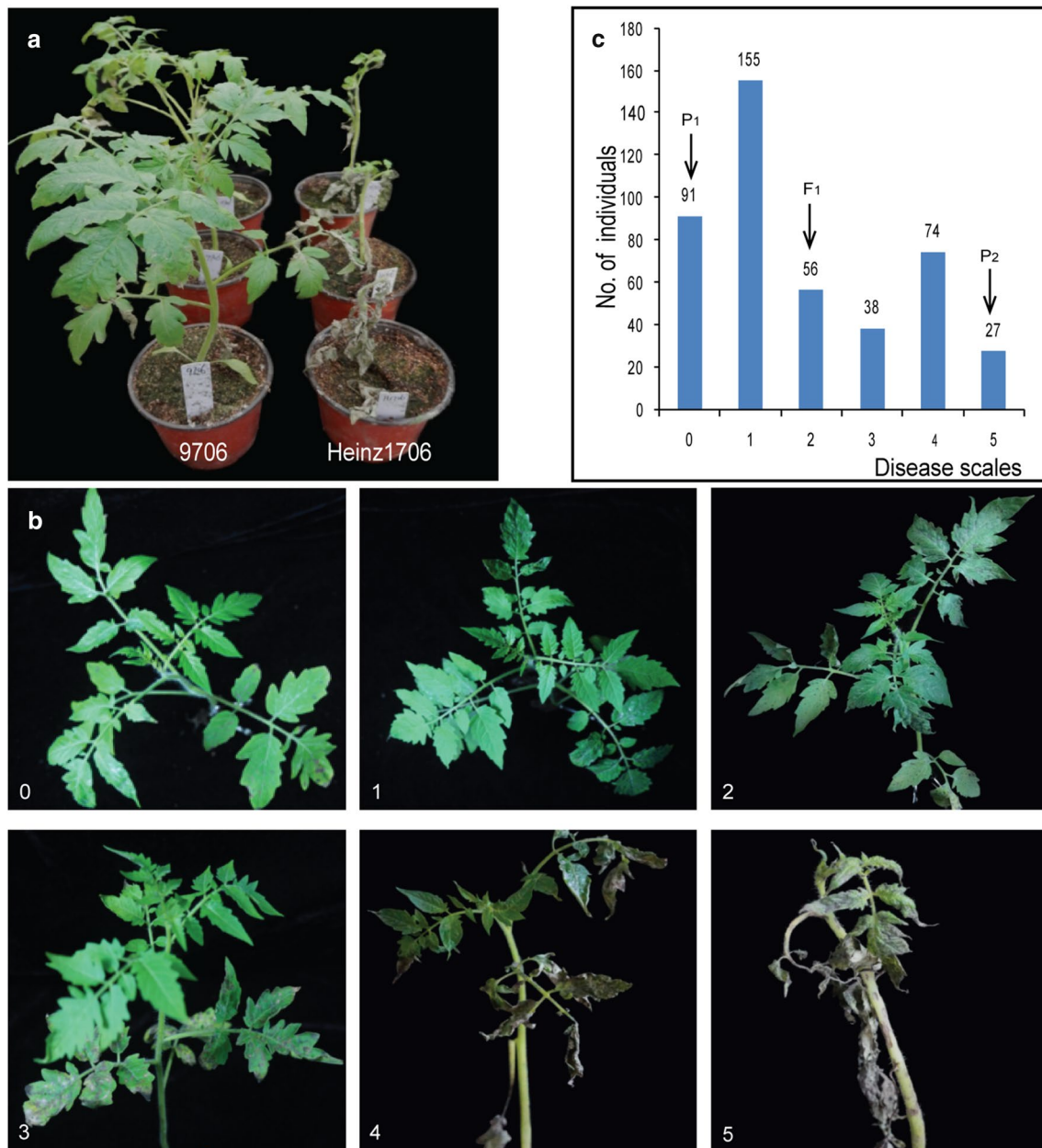


Fig. 1 Symptoms of parental lines (a) and F₂ population (b) after inoculation with *S. solani*. The number in each image indicates the symptom rating index (0–5). The frequency distribution of the disease scale in the F₂ population (c). The numbers above the bars indi-

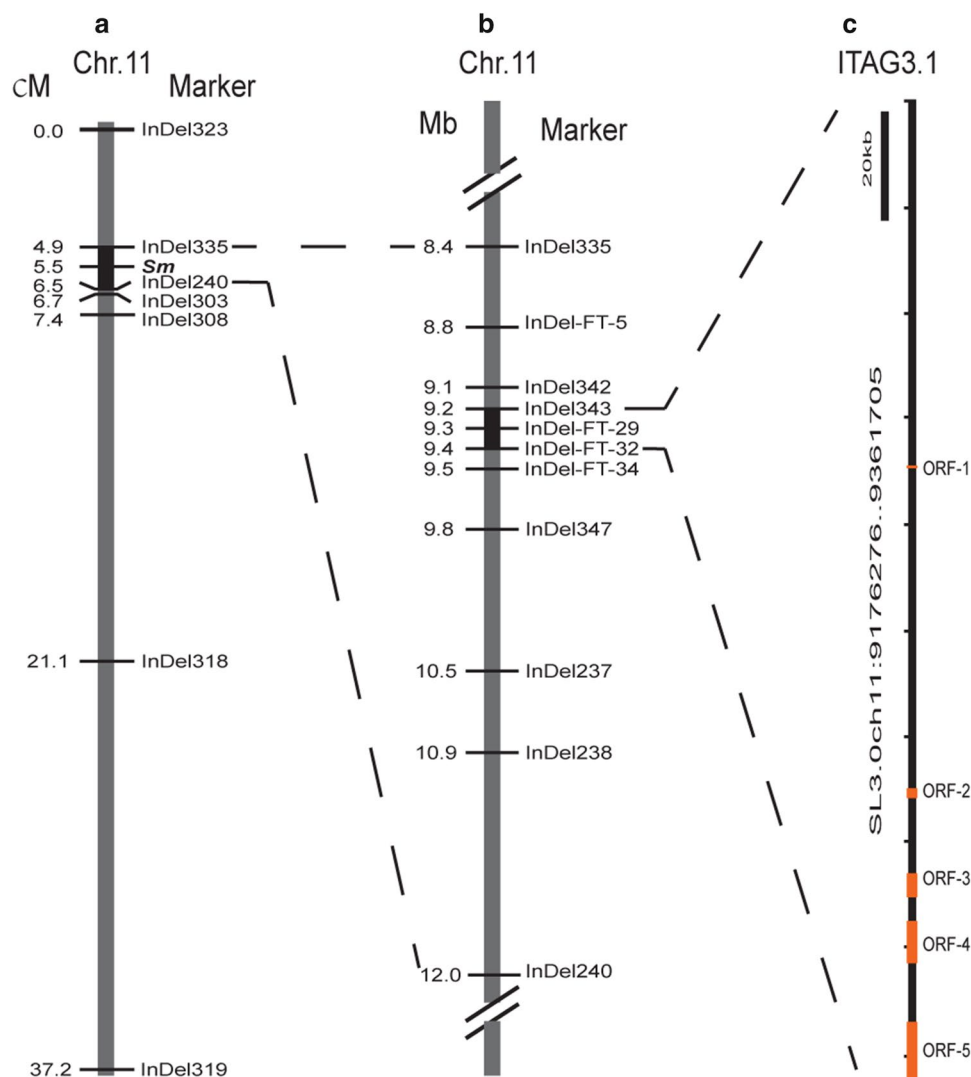
cate the number of individuals for each scale. The DS level of resistant parent “9706” (P₁)=0, susceptible parent “Heinz 1706” (P₂)=5, and F₁=1 or 2

Fine mapping of the *Sm* locus

Based on the phenotypic and genotypic data for the 441 F₂ individuals with seven markers, the *Sm* genetic map was created by JoinMap4.0 (Fig. 2a). All seven markers showed a linear order between their genetic and physical locations. Subsequently, another 4730 F₂ individuals were screened for recombinants using the flanking markers in batches. In brief, we first screened 2066 F₂ individuals with the markers

InDel335 and InDel240 and got 61 recombinants, analyzed these individuals with more markers between InDel335 and InDel240, narrowed down the mapping region. Then, we screened another F₂ individuals with new flanking markers and narrowed down to smaller region, and then the third batch. Eventually, the 15 selected F₂ recombinants were analyzed with nine new markers (Fig. 2b) and their F_{2,3} families were tested for gray leaf spot resistance. By combining the F₂ genotypes and the phenotypes of their corresponding

Fig. 2 Genetic and physical maps of the *Sm* gene and candidate genes analysis. **a** A genetic map of the *Sm* locus. Positions of the markers are indicated in cM. The linkage map was generated using 441 F₂ individuals using JionMap 4.0, **b** the physical map of the *Sm* locus. Positions of the markers are indicated in Mb. The *Sm* gene was located between markers InDel343 and InDel-FT-32 by recombinants and **c** the locations of the five candidate genes named as ORF-1, ORF-2, ORF-3, ORF-4 and ORF-5 at the *Sm* locus in the ‘Heinz 1706’ genome sequence



F_{2,3} families (Table 1), for example, the recombinant 7-81 was homozygous resistant for markers downstream of InDel343 and gray leaf spot resistance uniform in its F₃ family; thus, it was deduced that the *Sm* gene is downstream of InDel343. Similarly, from the recombinant 7-77, it was concluded that the *Sm* gene is upstream of InDel-FT-32. Therefore, the *Sm* was narrowed to an interval between markers InDel343 (SL3.0ch11_9176276) and InDel-FT-32 (SL3.0ch11_9361705), a 185-kb region on the ‘Heinz 1706’ genome that carries five predicted genes according to the tomato genome annotation (Fig. 2c).

GWAS analysis for the *Sm* locus

To confirm the map-based cloning results, we analyzed 289 diverse accessions using a GWAS for which a genome resequencing data set was available (Lin et al. 2014). We identified a strong association signal, $P = 1.12 \times 10^{-20}$, SNP: SL2.50ch11_9317565 (Fig. 3), for tomato gray leaf spot

resistance, which was coincident with the mapping analysis. Among the 289 accessions, most corresponded to the ordinary rule in which single-base ‘‘C’’ indicates susceptible and single-base ‘‘T’’ indicates resistance (Table S1), which represents a tight linkage between the SNP and the *Sm* gene.

Gene identification and expression

Five putative genes which named ORF-1 to ORF-5 were identified in the 185-kb region that corresponded to the *Sm* locus by searching the tomato genome annotation database in SGN (Fig. 2c; Table 2). However, none of them encoded a typical disease resistance protein. We performed a qRT-PCR analysis of the annotated genes in the leaf tissues of the two parental lines at 0, 2, 4 and 6 days post-inoculation, and the results are presented in Fig. S1. Among these five annotated genes, *Solyc11g018743* showed very low expression levels in both parents, which is consistent with the RNA sequencing coverage information in SGN. Among the other four

Table 1 Genotypes (F_2) and phenotypes (F_3) of recombinants

F_2 recombinant individual	Genotype (F_2) ¹										SNP ³	Phenotype (F_3 , progeny) ²	F_3 DS level	
	InDel 335	InDel FT-5	InDel 342	InDel 343	InDel FT-29	InDel FT-32	InDel FT-34	InDel 347	InDel 237	InDel 238				InDel 240
5-189	a	a	a	a	a	a	a	a	h	h	h	C	S	4, 5
6-277	a	a	a	a	a	a	a	h	h	h	h	C	S	4
6-25	b	b	b	b	b	b	b	h	h	h	h	T	R	0, 1
7-418	b	b	b	b	b	b	b	h	h	h	h	T	R	0
7-344	b	b	b	b	b	b	b	h	h	h	h	T	R	1, 2
7-77	b	b	b	b	b	b	b	h	h	h	h	T	R	0, 1, 2, 3
6-375	h	h	h	h	h	h	h	a	a	a	a	C/T	H	1, 3, 4
6-430	h	h	h	h	h	h	h	a	a	a	a	C/T	H	2, 3, 4
7-81	h	h	h	h	b	b	b	b	b	b	b	T	R	0, 1, 2
5-739	h	h	h	h	b	b	b	b	b	b	b	T	R	0, 1, 2
7-295	h	h	b	b	b	b	b	b	b	b	b	T	R	0, 1
7-514	b	b	h	h	h	h	h	h	h	h	h	C/T	H	2, 3, 4
7-56	a	a	h	h	h	h	h	h	h	h	h	C/T	H	2, 3, 4
7-91	h	h	a	a	a	a	a	a	a	a	a	C	S	4, 5
7-385	h	h	a	a	a	a	a	a	a	a	a	C	S	4, 5

¹a: homozygous for susceptible parent ‘Heinz 1706’, b: homozygous for resistant parent ‘9706’, h: heterozygous

²R, the F_3 family plants have a DS score of 0–3 and are considered as resistant. S, the F_3 family plants have a DS score of 4–5 and are regarded as susceptible. H, the F_3 family plants have a DS score of 0–5 and are regarded as heterozygous

³SNP means the association SNP in GWAS. The highlight in bold values the important recombinants that determine the *Srr* locus

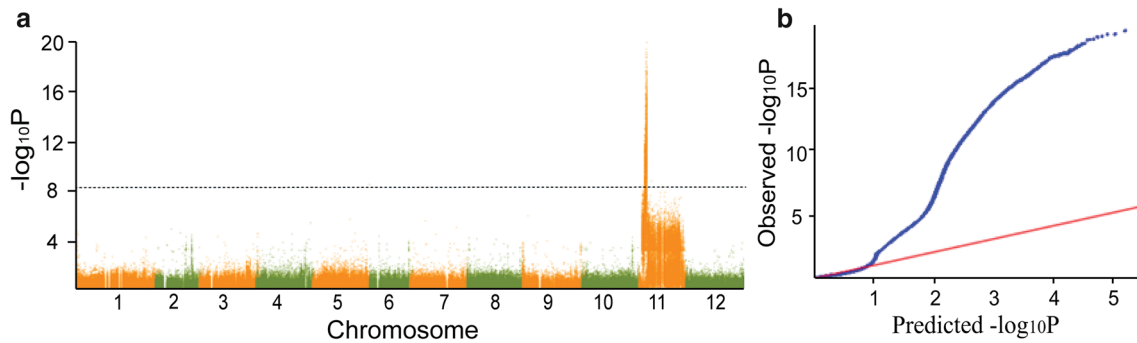


Fig. 3 Genome-wide association study of tomato gray leaf spot resistance. **a** Manhattan plot of the GWAS for tomato gray leaf spot resistance using the compressed mixed linear model (MLM). Negative \log_{10} -transformed P values from a genome-wide scan are plotted

against position on each of 12 chromosomes. Black horizontal dashed line indicates the genome-wide significance threshold. **b** Quantile–quantile plot of compressed MLM for tomato gray leaf spot resistance

Table 2 Annotated genes in the *Sm* region. List shows predicted genes in the 185-kb *Sm* region (ITAG3.1) in SGN (<https://solgenomics.net/>), and the primers used for the expression level analysis of the five genes

ORF number	SGN gene name	Position	Annotation	Primer for qRT-PCR	
ORF-1	<i>Solyc11g018710</i>	9,245,468..9,246,061 (+)	LOW QUALITY:DUF688 family protein	CTCCAAGATATGGTC AAAAAGC	TCCAAGAAAATGGC ACTC
ORF-2	<i>Solyc11g018715</i>	9,314,412..9,316,209 (-)	Retrovirus-related Pol polyprotein from transposon TNT 1-94	TTACTCTTCCACGAC ACCAA	AGGCCATGAGCATAT CACTA
ORF-3	<i>Solyc11g018740</i>	9,322,219..9,326,897 (+)	Inner membrane OXA1-like protein	TCACTACCGCCACCA CAAC	AGAAGGAACGGCAGA AGGA
ORF-4	<i>Solyc11g018743</i>	9,331,575..9,339,507 (+)	Membrane insertion protein, OxaA/YidC with tetratricopeptide repeat domain-containing protein	TCCTGGAAGTTTCAC CAAC	ATTGCGGAGCCAAAG AAGT
ORF-5	<i>Solyc11g018747</i>	9,350,344..9,361,393 (+)	Inner membrane ALBINO3-like protein	GGTTAATCCTGGAGA AATCA	TTGGCAATAGCAAGT CGTAG

genes, *Solyc11g018740* and *Solyc11g018747* showed lower expression levels in the susceptible genotype. In contrast, the expression of the *Solyc11g018715* gene in the susceptible plants was greater than in the resistant plants. What is more, the *Solyc11g018710* expression was not coincident on different time points; it is higher in susceptible plants than resistant before and 2 days after inoculation and antipodal at 4 days and 6 days post-inoculation. In conclusion, the putative genes showed small differences in expression levels in the two parents at some time points, but they were not coincident across all time points. Thus, further studies are required to ascertain a candidate gene for *Sm*.

Gene prediction and functional analysis

The 185-kb genomic DNA region in the Heinz1706 draft genome delimited by InDel343 and InDel-FT-32 were manually annotated. Thirty-three genes named ORF1 to ORF33 were predicted in this region by the FGENESH program,

which is more than the number of Heinz 1706 genome annotations. The details of these predicted genes are provided in supplemental Table S3.

To test the implication of the predicted genes on gray leaf spot resistance, specific VIGS constructs were designed to silence these predicted genes. Among the 33 predicted genes, 25 were amplified and assembled into TRV2 vector (Table S3). Subsequently, the infiltration mixture was vacuum-infiltrated into the germinating seeds of resistance line 9706 and the treated sprouts were sown. Photo-bleaching was observed on the true leaves of the PDS-silenced tomato seedling 20 days post-infection and sowing (Fig. 4a), indicating a successful silencing effect. Forty days later, plants with five fully expanded leaves were inoculated with *S. solani*. Those infiltrated with an empty TRV2 vector (Fig. 4b), water and most ORF-silenced plants showed no obvious symptoms except the ORF9-silenced plants. Among the 21 ORF9-silenced plants, eight showed leaf lesions, wilt and even

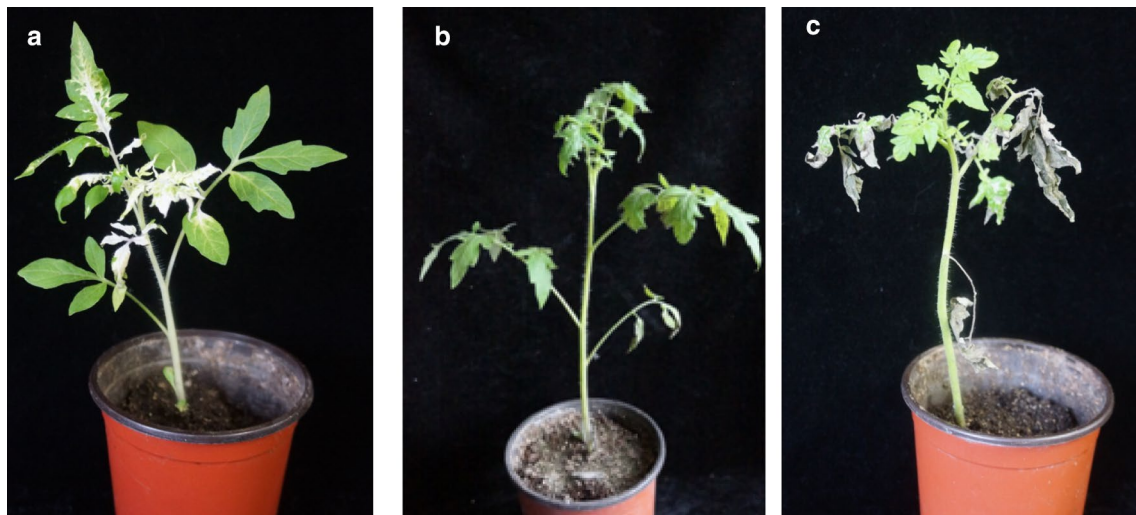


Fig. 4 The phenotype of plants inoculated with *S. solani* in virus induced gene silencing experiment. **a** Silencing of tomato *PDS* causes visible phenotypic symptoms, **b** control plant infiltrated with empty

TRV2 vector after inoculation with *S. solani*, **c** symptoms of ORF9-silenced tomato leaves after inoculation with *S. solani*

defoliation (Fig. 4c). The independent VIGS experiments in cotyledon infiltration method were performed, and the results were consistent. These results suggested that ORF9 was involved in *Sm*-mediated resistance and it seems to be good candidate of the *Sm* locus.

BAC library screening

According to the sequence and physical position of ORF9, we found unsequenced region with a large number of N in Heinz 1706 reference genome. In order to obtain the sequence covering the ORF9 region, an existing BAC was screened. Ultimately, a positive clone B80B15 was picked up and sequenced (Fig. S2). The whole BAC sequence of B80B15 carries an insert of 80,429 bp corresponding to an interval of 82,790 bp in the Heinz1706 reference genome starting from SL2.50ch11: 9198580 and ending at SL2.50ch11: 9281370. The ~ 80 kb BAC sequence was analyzed with online program FGENESH, and the result showed predicted gene harboring ORF9 consists of 14 exons of 5256 nucleotides, encoding a predicted polypeptide of 1751 amino acids. The predicted protein harbors Retrotrans_gag (Retrotransposon gag protein), zf-CCHC (Zinc knuckle), retropepsin_like (pepsin-like aspartate proteases), RVP_2 (Retroviral aspartyl protease), RT-LTR (reverse transcriptases), RVT_1 (reverse transcriptase) and RNase_HI_RT_Ty3 (Ty3/Gypsy family of RNase HI in long-term repeat retroelements) (Fig. S3). The conserved domain and structure characteristics demonstrated that the *Sm* candidate gene was not the typical type R gene.

Confirmation of linkage between marker Sm-InDel and the *Sm* gene

A 18-bp deletion was found downstream 15 kb from the *Sm* candidate gene when aligning the BAC sequence to Heinz 1706 reference genome and a co-dominant marker Sm-InDel was developed which produced a 122-bp fragment for resistant and a 140-bp fragment for susceptible alleles (Fig. 5). It was also tested on the 289 diverse accessions. The Sm-InDel marker predicted the gray leaf spot resistance of the germplasm, and consistent results between phenotype and genotype were obtained for 279 of the 289 accessions (Table S1), indicating a 96.5% accuracy rate.

Discussion

Sequencing technologies and applications of bioinformatics accelerate the genetic studies

Next-generation sequencing technologies are faster and more efficient and thus make it easier to establish associations between phenotypic traits and molecular markers or candidate genes, which are necessary for MAS in molecular breeding. With the rapid development of next-generation sequencing technologies and application of bioinformatics, many new strategies have been proposed to identify major quantitative trait loci or genes. GWAS has emerged a powerful approach for identifying genes underlying complex diseases at an unprecedented rate. Now, it has been applied to the dissection of complex traits in the crop plants rice

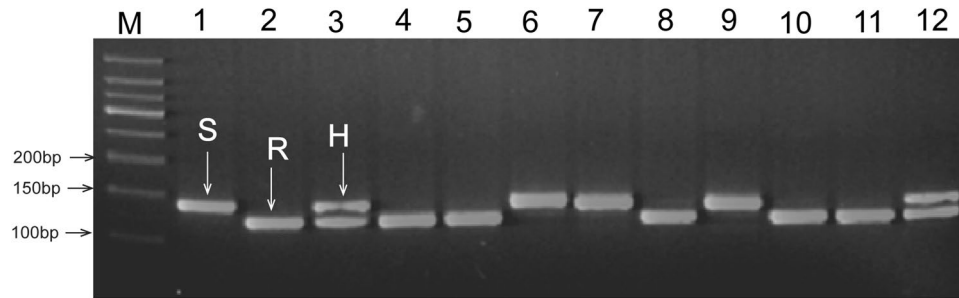


Fig. 5 PCR fragments amplified using marker Sm-InDel from tomato gray leaf spot resistant and susceptible genotypes resolved in a 3% agarose gel. Lane 1: Heinz 1706, Lane 2: 9706, Lane 3: F₁ (Heinz 1706×9706), Lane 4: LA2823, Lane 5: LA1589, Lane 6: Money-

maker, Lane 7: Ailsa Craig, Lane 8: M82, Lane 9: E-6203, Lane 10: Micro-Tom, Lane 11: 12 g-60, Lane 12: F₁ (Heinz 1706×LA1589). M, 50-bp DNA ladder; R, resistant; S, susceptible; H: heterozygote

(Huang et al. 2010, 2012; Wang et al. 2014), soybean (Zhou et al. 2015) and maize (Li et al. 2013; Tian et al. 2011). In this study, GWAS was carried out to identify the locus associated with gray leaf spot disease resistance in 289 tomato accessions, and a significantly associated locus was identified that was consistent with the map-based cloning analysis.

Gene prediction and functional analysis

The tomato genome is estimated to encode ~35,960 genes, majority of which are populated at distal euchromatic regions of the chromosomes (Khush and Rick 1968; Peterson et al. 1998; Van der Hoeven et al. 2002) with an approximate gene density of 6.7 kb/gene (Wang et al. 2005). Based on the ‘Heinz 1706’ genome annotations, there were five annotated genes in the defined region, which is less than expected. Moreover, these annotated genes expression levels were not significantly different in the two parents across all of the time points. By analyzing the target region in the ‘Heinz 1706’ reference genome, we found several regions that were not sequenced in the corresponding 185-kb region, which might be one reason for the scarcity of annotated genes. Thus, we predicted genes using FGENESH within our mapping region, and 33 putative open reading frames were found in the corresponding 185-kb region, which included the five annotated genes. Subsequently, the virus-induced gene silencing approach was performed for the 33 putative open reading frames to determine the candidate genes. Due to such reasons as low level of putative gene expression, mistakes on gene prediction, unsatisfactory primer design, eventually, 25 VIGS vectors were constructed and used for gene functional analysis. In this study, we performed the VIGS experiments using SVI method, which was a simple, effective and low-cost agroinoculation method to elicit VIGS in tomato. Ultimately, the ORF9-silenced plants showed susceptible to *S. solani*. This result indicated ORF9 was involved in *Sm*-mediated resistance, and it seems to be good candidate of the *Sm* locus.

VIGS is a transcript suppression technique used for gene function analysis. This technique offers a tremendous advantage for knocking down the target genes due to the needless for transformation in plant (Baulcombe 1999; Dinesh-Kumar et al. 2003). Agroinoculation is a widely used method to deliver VIGS vectors into plants for RNA silencing (Liu et al. 2002; Gossele et al. 2002). However, the current agroinfiltration methods have their limitations. Among the existing methods, leaf infiltration method is widely used to introduce the infiltration solution into seedling leaves using a needle-less syringe, vacuum machine, or high-pressure spray gun, but they are laborious for large-scale VIGS screens. However, the barrier has been overcome using the SVI method which proved to be applicable for high-throughput VIGS screens. With the improvement in the tomato genome sequence and annotation, a large number of candidate genes of interest could be examined at the levels of transcriptome in present. Therefore, SVI will be a better agroinoculation method to implement large-scale VIGS experiments with minimal manpower in a relatively short time.

Candidate gene cloning through BAC sequencing

Based on the analysis of the ORF9 sequence and its physical position in Heinz1706 reference genome, we found two GAPs before and after this target region and we tried to sequence the GAPs using the PCR products. Unfortunately, the full length of candidate gene could not be amplified from the *Sm*-containing tomato lines with the primers designed according to the Heinz1706 reference genome. In order to obtain the sequence covering the ORF9 region, a BAC library was screened and the positive clone was sequenced. The result demonstrated that the *Sm* candidate gene does not belong to the typical type R gene which contains domains such as NBS-LRR. In plant, there also exists the particular, such as *Ty1*, which was a completely new class gene code for an RNA-dependent RNA polymerase and increased cytosine methylation of viral genomes to confer resistance

(Verlaan et al. 2013; Butterbach et al. 2014). The *Mlo* gene in barley, whose resistance is caused by rapid formation of large cell wall appositions at the encounter sites, prevents penetration by the fungus (Jørgensen 1992). Anyway, genetic transformation is required to identify the candidate gene in the further study.

Applicability of markers in MAS

Marker-assisted selection (MAS) is a powerful tool in plant breeding. The development of molecular markers tightly linked with the target gene is a prerequisite for efficient MAS. Until now, there have been few reports of molecular research concerning tomato resistance to gray leaf spot. An incipient study mapped *Sm* on chromosome 11 between RFLP markers TG110 and T10, due to the distance to the target gene and not PCR-based, these two markers are not breeder friendly. In 2009, Ji and Scott reported a recessive CAPS marker CT55, which cannot distinguish between heterozygous and homozygous susceptible plants. Therefore, its use may be partly restricted in practice. In a recent research, *Sm* gene was mapped and screened based on bulked segregant analysis in combination with genome resequencing, accompanied with a SCAR marker D5 designed (Yang et al. 2017). However, this marker hardly distinguishes resistant and susceptible tomato germplasm in our verification experiment, which indicated that the marker showed poor applicability in tomato breeding. In the present study, a co-dominant marker Sm-InDel was developed that produced a 122-bp or a 140-bp fragment for resistant or susceptible alleles, respectively. It can detect gray leaf spot resistance in germplasm with 96.5% accuracy and easily resolved in a standard agarose gel. The inconsistency between phenotypes and genotypes in 289 diverse accessions might result from the following: (1) a deficiency in the inoculation tests; (2) crossing-over occurring. This result indicates that the co-segregating marker Sm-InDel is very close to the gene *Sm* and could be used in MAS breeding for gray leaf spot resistance.

Author contribution statement XMS performed the experiments, analyzed the data, and wrote the manuscript; GTZ analyzed the data; ZJH, XXW and YMG revised the manuscript; BJL provided pathogen and inoculation method; YCD, WCY and JCG conceived and designed the experiments, and revised the manuscript.

Acknowledgements This work was supported by the National Key Research and Development Program of China (2016YFD0101703), the National Natural Science Foundation of China (31471874) and the Key Laboratory of Horticultural Crops Genetic Improvement, Ministry of Agriculture of the People's Republic of China.

Compliance with ethical standards

Conflict of interest The authors of this manuscript declare that they have no conflict of interest.

References

- Andrus CF, Reynard GB, Wade BL (1942) Relative resistance of tomato varieties, selections, and crosses to defoliation by *Alternaria* and *Stemphylium*. US Department of Agriculture Circular 652
- Bashi E, Pilowsky M, Rotem J (1973) Resistance in tomatoes to *Stemphylium floridanum* and *Stemphylium botryosum* f. sp. *lycopersici*. *Phytopathology* 63:1542–1544
- Baulcombe DC (1999) Fast forward genetics based on virus-induced gene silencing. *Curr Opin Plant Biol* 2:109–113
- Behare J, Laterrot H, Sarfatti M, Zamir D (1991) Restriction fragment length polymorphism mapping of the *Stemphylium* resistance gene in tomato. *Mol Plant-Microbe Interact* 4:489–492
- Blancard D, Laterrot H, Moretti A, Pichot M (1986) Les *Stemphylium* rencontrés sur tomate. *Phytopathol Mediterr* 25:140–144
- Butterbach P, Verlaan MG, Dulleman A, Lohuis D, Visser RGF, Bai Y, Kormelink R (2014) Tomato yellow leaf curl virus resistance by *Ty-1* involves increased cytosine methylation of viral genomes and is compromised by cucumber mosaic virus infection. *Proc Natl Acad Sci* 111:12942–12947
- Cedeño L, Carrero C (1997) First report of tomato gray leaf spot caused by *Stemphylium solani* in the Andes region of Venezuela. *Plant Dis* 81:1332
- Chai AL, Du GF, Shi YX, Xie XW, Li BJ (2015) Leaf spot on sweet potato (*Ipomoea batatas*) caused by *Stemphylium solani*, a new disease in China. *J Phytopathol* 163:1046–1049
- Chen X (1996) The tomato disease resistance identification technology of France. *World Agric* 5:010 (in Chinese)
- Dennett RK (1950) The association of resistance to *Fusarium* wilt and *Stemphylium* leaf spot in Tomato, *Lycopersicon esculentum*. *Proc Am Soc Hortic Sci* 56:353–357
- Dinesh-Kumar SP, Anandalakshmi R, Marathe R, Schiff M, Liu Y (2003) Virus-induced gene silencing. *Methods Mol Biol* 236:287–293
- Everts K, Armentrout DK (2001) Report of leaf spot of spinach caused by *Stemphylium botryosum* in Maryland and Delaware. *Plant Dis* 85:1209
- Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol Biol Rep* 13:207
- Gossele V, Fache I, Meulewaeter F, Cornelissen M, Metzclaff M (2002) SVISS—a novel transient gene silencing system for gene function discovery and validation in tobacco plants. *Plant J* 32:859–866
- Hendrix JW, Frazier WA (1949) Studies on the inheritance of *Stemphylium* resistance in tomatoes. Hawaii Agricultural Experiment Station, technical bulletin no. 8
- Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961–967
- Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 44:32–39
- Ji Y, Scott JW (2009) A CAPS marker linked to the tomato gray leaf spot (*Stemphylium* sp.) resistance gene *Sm*. *TGC Rep* 59:29–31
- Jørgensen IH (1992) Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. *Euphytica* 63:141–152

- Kang HM, Su JH, Service SK, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354
- Khush GS, Rick CM (1968) Cytogenetic analysis of the tomato genome by means of induced deficiencies. *Chromosoma* 23:452–484
- Kim BS, Cho HJ, Hwang HS, Cha YS (1999) Gray leaf spot of tomato caused by *Stemphylium solani*. *Plant Pathol J* 15:348–350
- Kim BS, Yu SH, Cho HJ, Hwang HS (2004) Gray leaf spot in peppers caused by *Stemphylium solani* and *S. lycopersici*. *Plant Pathol J* 20:85–91
- Koike ST, Henderson DM, Butler EE (2001) Leaf spot disease of spinach in California caused by *Stemphylium botryosum*. *Plant Dis* 85:126–130
- Krattinger SG (2009) Molecular characterization of the durable broad-spectrum disease resistance gene *Lr34* in wheat (*Triticum aestivum* L.). Universität Zürich, Zürich
- Li B, Zhou Y, Li J, Xie X (2010) Diagnosis of diseases By Dr Li Bao Ju: the diagnosis and prevention of tomato *Stemphylium* leaf spot disease (gray leaf spot). *China Veg* 2010:24–25 (in Chinese)
- Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet* 45:43–50
- Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, Zhang Z, Lun Y, Li S, Wang X, Huang Z, Li J, Zhang C, Wang T, Zhang Y, Wang A, Zhang Y, Lin K, Li C, Xiong G, Xue Y, Mazzucato A, Causse M, Fei Z, Giovannoni JJ, Chetelat RT, Zamir D, Stadler T, Li J, Ye Z, Du Y, Huang S (2014) Genomic analyses provide insights into the history of tomato breeding. *Nat Genet* 46:1220–1226
- Liu Y, Schiff M, Dinesh-Kumar SP (2002) Virus-induced gene silencing in tomato. *Plant J* 31:777–786
- Liu A, Sun J, Tao X, Qu Q, Xu D, Wang T (2004) The occurrence and comprehensive prevention of tomato gray leaf spot in protect area. *China Plant Prot* 24:23–24 (in Chinese)
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25(4):402–408
- Min J, Kim B, Cho K, Yu S (1995) Grey leaf spot caused by *Stemphylium lycopersici* on tomato plants. *Korean J Plant Pathol* 11:282–284
- Nasehi A, Kadir JB, Zainal Abidin MA, Wong MY, Abed Ashtiani F (2012a) First report of gray leaf spot on pepper caused by *Stemphylium solani* in Malaysia. *Plant Dis* 96:1227
- Nasehi A, Kadir JB, Zainal Abidin MZ, Wong MY, Mahmodi F (2012b) First report of tomato gray leaf spot disease caused by *Stemphylium solani* in Malaysia. *Plant Dis* 96:1226
- Nasehi A, Kadir JB, Esfahani MN, Mahmodi F, Ghadirian H, Abed Ashtiani F, Golkhandan E (2013) An outbreak of leaf spot caused by *Stemphylium solani* on eggplant in Malaysia. *Plant Dis* 97:689
- Peterson DG, Pearson WR, Stack SM (1998) Characterization of the tomato (*Lycopersicon esculentum*) genome using in vitro and in situ DNA reassociation. *Genome* 41:346–356
- Rotem J, Cohen Y, Wahl I (1966) A new tomato foliage disease in Israel caused by *Stemphylium botryosum*. *Can J Plant Sci* 46:265–270
- Spielmeier W, Sharp PJ, Lagudah ES (2003) Identification and validation of markers linked to broad-spectrum stem rust resistance gene *Sr2* in wheat (*Triticum aestivum* L.). *Crop Sci* 43:333–336
- Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* 43:159–162
- Vakalounakis DJ, Markakis EA (2013) First report of *Stemphylium solani* as the causal agent of a leaf spot on greenhouse cucumber. *Plant Dis* 97:287
- Van der Hoeven R, Ronning C, Giovannoni J, Martin G, Tanksley S (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell* 14:1441–1456
- Van Ooijen J (2006) JoinMap® 4: software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen* 33:10.1371
- Verlaan MG, Hutton SF, Ibrahim RM, Kormelink R, Visser RGF, Scott JW, Edwards JD, Bai Y (2013) The tomato yellow leaf curl virus resistance genes *Ty-1* and *Ty-3* are allelic and code for DFDGD-class RNA-dependent RNA polymerases. *PLoS Genet* 9:e1003399
- Wang H, Qi M, Cutler AJ (1993) A simple method of preparing plant samples for PCR. *Nucl Acid Res* 21:4153
- Wang Y, Van Der Hoeven RS, Nielsen R, Mueller LA, Tanksley SD (2005) Characteristics of the tomato nuclear genome as determined by sequencing undermethylated *EcoRI* digested fragments. *Theor App Genet* 112:72–84
- Wang C, Yang Y, Yuan X, Xu Q, Feng Y, Yu H, Wang Y, Wei X (2014) Genome-wide association study of blast resistance in indica rice. *BMC Plant Biol* 14:1
- Xie X (2017) The etiology of *stemphylium* leaf spot on vegetables and its resistance to boscalid in China. Ph.D. dissertation (in Chinese)
- Yan H, Fu D, Zhu B, Liu H, Shen X, Luo Y (2012) Sprout vacuum-infiltration: a simple and efficient agroinoculation method for virus-induced gene silencing in diverse *solanaceous* species. *Plant Cell Rep* 31:1713–1722
- Yang H, Zhao T, Jiang J, Wang S, Wang A, Li J, Xu X (2017) Mapping and screening of the tomato *Stemphylium lycopersici* resistance gene, *Sm*, based on bulked segregant analysis in combination with genome resequencing. *BMC Plant Biol* 17:266
- Zheng L, Huang J, Hsiang T (2008) First report of leaf blight of garlic (*Allium sativum*) caused by *Stemphylium solani*, in China. *Plant Pathol* 57:380
- Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol* 33:408–414