Mol Breeding (2020) 40: 45
https://doi.org/10.1007/s11032-020-01126-8

Fine mapping of QTLs for resistance to *Phytophthora nicotianae* in flue-cured tobacco using a high-density genetic map

Daping Gong · Mingli Chen · Yang Sun · Yuqin Zhang · Xingtan Zhang · Xiuhong Xu

Received: 5 September 2019 / Accepted: 3 April 2020 / Published online: 25 April 2020
© Springer Nature B.V. 2020

Abstract Tobacco (*Nicotiana tabacum* L.) is an important agronomic crop and model system for studies of plant-pathogen interactions. Black shank, caused by *Phytophthora nicotianae*, is an important disease affecting tobacco production worldwide. In this study, a mapping population of 177 $F_{7:8\text{-}9}$ recombinant inbred lines was generated from a cross between the highly resistant cultivar 'Yunyan 85' and a susceptible line 'Dabaijin 599'. A high-density genetic linkage map containing 7734 single-nucleotide polymorphic markers based on restriction site-associated DNA tag sequencing technology was used to finely map quantitative trait loci (QTL) for resistance to *P. nicotianae*. A total of 10 QTLs were detected as being associated with resistance to *P. nicotianae* across multiple environments, and two major QTL *qBS7* and *qBS14* were repeatedly identified under all five environments. They explained 16.48-62.20% and 3.94-11.29% of the phenotypic variance with high LOD score, respectively. One hundred thirty-eight candidate genes were identified for two major QTLs *qBS7* and *qBS14*, and annotation analysis showing that several predicted genes encoded proteins associated with plant defense response to pathogens. This high-density single-nucleotide polymorphic genetic linkage map of flue-cured tobacco based on restriction site-associated DNA sequencing was useful in the QTL finely mapping of resistance to *P. nicotianae*. This study increases our understanding of the genetics of resistance to *P. nicotianae* and aids in marker-assisted selection.

Keywords Flue-cured tobacco · *Phytophthora nicotianae* · Quantitative trait loci · Single-nucleotide polymorphic · Genetic map

Abbreviations

QTL	Quantitative trait loci
RIL	Recombinant inbred line
SNP	Single nucleotide polymorphism
NBS-LRR	Nucleotide-binding site-leucine-rich repeat
LOD	Log-likelihood

Daping Gong and Mingli Chen contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11032-020-01126-8) contains supplementary material, which is available to authorized users.

D. Gong · M. Chen · Y. Sun · Y. Zhang · X. Xu (✉)
Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao 266101, China
e-mail: xuxiuhong@caas.cn

Y. Sun · Y. Zhang
Graduate School of the Chinese Academy of Agricultural Sciences, Beijing 100081, China

X. Zhang
Fujian Agriculture and Forestry University, Fuzhou 350002, China

Introduction

The *Phytophthora* genus contains some of the most destructive plant pathogens that attack a wide range of

economically important plant species worldwide. Black shank, caused by the soil-borne pathogen *Phytophthora nicotianae*, is the most important disease affecting tobacco (*Nicotiana tabacum* L.) production in many countries. In contrast to chemical control strategies, genetic resistance is the most cost-effective protective method to reduce economic losses caused by this pathogen.

Monogenic-specific resistance to race 0 of *P. nicotianae*, which is conferred by the *Php* and *Phl* genes, has been introgressed into *N. tabacum* from *Nicotiana plumbaginifolia* (Apple 1962; Chaplin 1962) and *Nicotiana longiflora* (Valleau et al. 1960), respectively. However, the widespread planting of cultivars possessing the *Php* or *Phl* gene has resulted in race shifts from race 0 to race 1. Race 1 of *P. nicotianae* is able to overcome resistance, which is becoming prevalent with continuous planting (Sullivan et al. 2005). Additionally, race 2 of *P. nicotianae* has been reported in South Africa, and race 3 is virulent on varieties containing the *Phl* gene (van Jaarsveld et al. 2002; Gallup and Shew 2010; McIntyre and Taylor 1978). High levels of resistance to both race 0 and race 1 of *P. nicotianae* are associated with an introgressed *Nicotiana rustica* genomic region, designated as *Wz* (Drake et al. 2015; Drake and Lewis 2013). Subsequently, *Wz*-adapted isolates of *P. nicotianae* have also been found (McCorkle et al. 2018). Because of the low durability of monogenic resistance mechanisms, a rotation of cultivars with single-gene resistance and cultivars with a high level of polygenic resistance would reduce disease incidence and minimize pathogen race shifts (Sullivan et al. 2005).

Two primary polygenic black shank-resistance sources are from the cigar tobacco 'Florida 301' (Tisdale 1931) and 'Beinhart-1000' (Heggestad and Lautz 1957; Chaplin 1966; Wills 1971; Nielsen 1992). The 'Beinhart-1000' cultivar expresses a high level of partial resistance to both race 0 and race 1 of *P. nicotianae* (Chaplin 1966). Vontimitta and Lewis (2012a and b) detected two major quantitative trait loci (QTLs) associated with resistance using a doubled haploid mapping population generated from a 'Beinhart 1000' × 'Hicks' cross under field and growth chamber environmental conditions (Vontimitta and Lewis 2012a; Vontimitta and Lewis 2012b). The largest QTL on linkage group (LG) 8 and the second QTL on LG 4 explained 25.4-54.7% and 16.8-20.4% of the phenotypic variance, respectively. Because of the associations between resistance and undesirable cigar-type characteristics, the resistance of this line is less utilized in burley and flue-cured tobacco cultivars (Nielsen 1992).

Based on pedigree information, most varieties of flue-cured tobacco are believed to have derived polygenic black shank resistance from cigar tobacco 'Florida 301'. 'Florida 301' exhibits partial resistance in nature and is non-race-specific (Tisdale 1931). However, this type of resistance appears to be negatively correlated with yield in flue-cured tobacco. DNA markers closely associated with major genes affecting resistance might increase the possibility of transferring this resistance to commercial cultivars without the negative influence on yield. Using microsatellite markers, Xiao et al. (2013) identified 11 QTLs associated with a high level of black shank resistance in the cigar tobacco 'Florida 301' (Xiao et al. 2013). Among those, the largest effect QTL explained 16.9-18.6% of the phenotypic variation. The QTL also had the greatest phenotypic effects in a 'Beinhart 1000' × 'Hicks' doubled haploid (DH) population (Vontimitta and Lewis 2012b) and a 'K346' × 'TI1068' recombinant inbred line (RIL) population (Drake-Stowe et al. 2017). Recently, the QTL has been localized to within a genetic interval of approximately 3 cM using SNP markers and near-isogenic lines (NILs) (Ma et al. 2019).

A high-density simple sequence repeat (SSR)-based linkage map was generated containing 2318 microsatellite markers in tobacco (Bindler et al. 2011), and the vast majority of QTL mapping studies in *N. tabacum* have used these microsatellite markers (Drake-Stowe et al. 2017; Xiao et al. 2013; Vontimitta and Lewis 2012a; Vontimitta and Lewis 2012b). Owing to the low polymorphism rates revealed among tobacco lines, especial flue-cured lines, the efficiency and accuracy of QTL mapping using microsatellite markers were limited. The rapid development of next-generation sequencing technologies has been used to discover single-nucleotide polymorphisms (SNPs) and complete tobacco genome sequencing including four allotetraploid *Nicotiana* (*N. benthamiana* and three *N. tabacum* cultivars, TN90, K326, and *Basma xanthi*) and three diploid *Nicotiana* (*N. otophora*, *N. sylvestris*, and *N. tomentosiformis*) (Bombarely et al. 2012; Edwards et al. 2017; Sierro et al. 2013; Sierro et al. 2014). The first SNP linkage map for *N. tabacum* was recently published based on an F_2 population using specific length amplified fragment sequencing (SLAF-seq), and two major QTLs were identified for curing trait (Gong et al. 2016).

Although a number of QTLs have been identified for black shank resistance, few QTLs have been finely mapped across genetic backgrounds and/or environments. In the current study, we aimed to improve the efficiency and accuracy of black shank resistance-related QTL mapping in tobacco using a high-density genetic linkage map, comprising 7734 SNP markers, through restriction site-associated DNA tag sequencing (RAD-seq) technology on 177 $F_{7:8\text{-}9}$ RILs. The RIL population was generated from a cross between highly resistant flue-cured tobacco 'Yunyan 85' and susceptible flue-cured tobacco 'Dabaijin 599'. We finely mapped QTLs for black shank resistance under four field and one greenhouse conditions.

Materials and methods

Plant materials and DNA extraction

An RIL mapping population containing 177 $F_{7:8\text{-}9}$ progenies was developed from a cross between two flue-cured tobacco varieties with distinctly different resistance levels to *P. nicotianae*. 'Yunyan 85' and 'Dabaijin 599' are considered to be the highly resistant and susceptible parents, respectively. 'Yunyan 85' was developed from the cross 'Yunyan 2' × 'K326', and 'Dabaijin 599' is a Chinese tobacco landrace (Tan et al. 1997). Young leaves were collected from parents and individual plants of the RIL population and frozen in liquid nitrogen for DNA extraction. Total genomic DNA was extracted using the cetyl trimethylammonium bromide method. DNA qualities were evaluated using NanoDrop 2000 spectrophotometer (Thermo Scientific, Milan, Italy) and electrophoresis in agarose gels.

Disease evaluation

Disease evaluation of all 177 RILs, and their parents was conducted in four fields and one greenhouse.

Black shank disease nurseries were going at the field station in Shandong, Zhucheng (N 35° 59′, E 119° 24′) in 2017 (ZC2017), 2018 (ZC2018), and 2019 (ZC2019), and in Shandong, Jimo (N 36° 38′, E 120° 45′) in 2018 (JM2018). Plants were transplant in single 12-plant row plots following a randomized complete block design. Two replications were completed in ZC2017 and JM2018, and with three replications in ZC2019 and with four replications in ZC2018. The susceptible cultivar 'Dabaijin 599' was planted in every third plot to check the uniformity of disease pressure. The inter-row spacing was 1.2 m and the within-row plant spacing was 0.5 m at both locations. Starting 45 days after transplanting, the severity of plants killed by black shank were evaluated at 15-day intervals using an empirical 6-point scale (GB/T23222-2008, China), where 0 = no symptoms; 1 = up to 33% stem lesions or chlorotic leaves; 3 = up to 50% stem lesions or chlorotic leaves; and 5 = up to 90% stem lesions or 66% chlorotic leaves; 7 = up to 100% stem lesions or chlorotic leaves; 9 = death. Disease index (DI) scores were calculated using the following formula: $DI = 100 \times \sum$[(disease evaluation scale score × number of plants with each scale score)]/(total number of plants observed × the highest disease evaluation scale score)].

The RIL population and parents were also evaluated for black shank resistance in the greenhouse in Shandong, Jimo, in 2019 (GH2019) according to Vontimitta et al. (Vontimitta and Lewis 2012a). Isolates of *P. nicotianae* race 0 were provided by Tobacco Research Institute of Chinese Academy of Agricultural Sciences Pathology Department. The experimental design was a randomized complete block design with three replications. Experimental units consisted of 8 plants contained within an 8-compartment segment of a nursery tray. Approximately 50 days after germination, plants were inoculated by wounding the stem base with a knife and inserting four *P. nicotianae*-infested oat grains into the soil. The temperature was maintained at 30 °C and soil moisture was uniformly maintained by subirrigation. Approximately 7 days after inoculation, the disease severity of plants was recorded at 7-day intervals. Five data from four fields and one greenhouse environments were used for analysis of variance (ANOVA) using SPSS 23.0 software.

RAD library preparation and sequencing

The RAD library was constructed following the protocol described by Baird et al. (2008), with minor modifications. Briefly, genomic DNA from the parents and each of the 177 progenies were digested for 15 min at 37 °C in a 50-μL reaction containing 20 units (U) of *Eco*RI (New England Biolabs, USA). A modified Illumina P1 adapter (USA) was added to the samples. Sample were pooled and randomly sheared to an average size of 500 bp. DNA fragments of 300-500 bp were isolated using the QIAquick Gel Extraction Kit (Qiagen,

Germany). The ends of the DNA were repaired using the Quick Blunting kit Enzyme Mix (New England Biolabs). Then, a modified Solexa P2 adapter was ligated to the DNA fragments (Illumina). Finally, purified and eluted DNA products were PCR-amplified using the Phusion Master Mix (New England Biolabs). The prepared DNA libraries were sequenced using Illumina HiSeq X Ten instrumentation.

SNP identification and genotyping

To ensure high-quality genotype calling, raw RAD sequence reads of low quality (quality score ≤ Q20) and without unique barcodes were filtered. After the barcodes and the terminal bases were trimmed, the clean reads from each individual were mapped onto the tobacco reference genome (Edwards et al. 2017) using BWA software (Li and Durbin 2009). Input data was prepared by SAMtools, and SNP calling was performed by GATK (Li et al. 2009; DePristo et al. 2011). SNP loci with more than three continuous SNPs in 300 bp were filtered out first. Because one SNP could contain at most four genotypes for diploid species, SNP loci with more than four alleles were discarded. All SNP loci were genotyped for consistency at the parental and offspring SNP loci. Before genetic map construction, all of the markers were filtered using the following criteria: First, sequence depths were > 10-fold in the parents. Second, markers with more than 30% missing data in the RIL population were filtered. Third, the chi-squared test was performed to examine the segregation distortion. Markers with significant segregation distortions ($P < 0.01$) were initially excluded from the map construction and were then added as accessory markers. Finally, markers having parental homozygous genotype (aa × bb) were used to construct the genetic map.

Linkage map construction and analysis

The HighMap strategy was utilized to order the SNP markers and correct genotyping errors (Liu et al. 2014). The enhanced algorithm of Gibbs sampling, spatial sampling, and simulated annealing was used to order markers (Van Ooijen 2011; Jansen et al. 2001). The SMOOTH algorithm was used to correct genotyping errors (van Os et al. 2005). The missing genotypes in progenies were imputed by the *k*-nearest neighbor algorithm (Huang et al. 2011). The skewed markers were added to the linkage map using a multipoint method of maximum likelihood. The map distances were calculated using the Kosambi mapping function (Kosambi 1944). Heat and haplotype maps were constructed to evaluate the map quality (West et al. 2006). All of the sequences of the SNP markers in the linkage map were aligned back to the physical sequences of the reference genome and two ancestral parents, *N. sylvestris* and *N. tomentosiformis*, using a Basic Local Alignment Search Tool.

QTL detection and candidate gene identification

QTL IciMapping 4.1 (Meng et al. 2015) was used to identify QTLs by the inclusive composite interval mapping method (Li et al. 2007). The threshold of the log-likelihood (LOD) scores was more than 2 and determined using 1000 permutations for $P = 0.05$. Additive and phenotypic contribution rates were estimated. When the positions of the QTLs overlapped within a range of 10 cM, they were interpreted to be the same QTL. The major QTLs, detected in five environments, were selected to identify the candidate genes. Based on the positions of the flanking markers, all of the genes within the confidence interval were identified as candidate genes.

Results

Genetic analysis of resistance to *P. nicotianae*

'Yunyan 85' exhibited a very high level of resistance in five environments, and with disease indices (DIs) less than 20. 'Dabaijin 599' exhibited a high level of susceptible reactions, with DIs of above 90. The RIL population of 177 lines displayed a wide range of variation, with DIs ranging from 0 to 100 and a mean DI value of 55 for the four field environments (Fig. 1) and one greenhouse environments (Fig. 2). Few lines exhibited DIs that were greater than those of 'Dabaijin 599' and lower than those of 'Yunyan 85'. The RIL population displayed significantly skewed distributions in the resistance direction, indicating that polygenes, including major genes, control resistance to black shank. The DI data for RILs in five environments were subjected to an ANOVA, which detected highly significant differences among lines ($P < 0.0001$) and significant environment × line interactions ($P < 0.0001$) (Table 1).

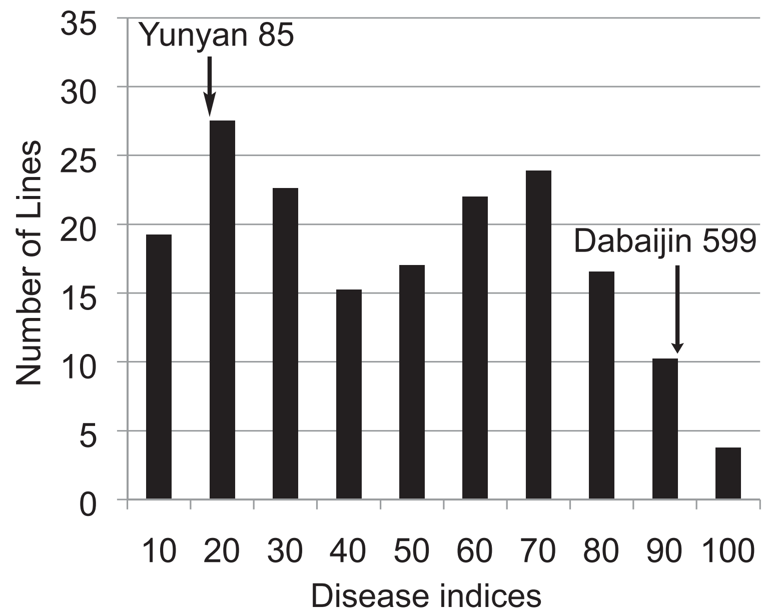

Fig. 1 Frequency distribution for the RIL population exhibiting disease indices averaged over four field environments. The disease indices for parental lines are indicated by arrows

SNP genotyping

A total of 179 RAD-seq libraries, two parental and 177 progenies, were constructed and sequenced in Illumina HiSeq X Ten. After removing low quality sequences, 4,618,576,318 high-quality paired-end reads were obtained, with each read being ~151 bp in length, representing ~1394.81 GB of sequences. The clean reads were deposited at NCBI under the bioproject. Of the clean reads, 34,944,329 were from 'Yunyan 85' and 39,685,656 were from 'Dabaijin 599', with the remaining being from the 177 RILs (Fig. 3). Read numbers for the individuals ranged from 10,648,452 to 68,786,148, with an average of 25,672,013. The filtered reads were then mapped to the reference tobacco genome. The average depths of sequencing were 7.95-fold in 'Yunyan 85', 8.40-fold in 'Dabaijin 599', and 5.94-fold in the progenies (Fig. 3).

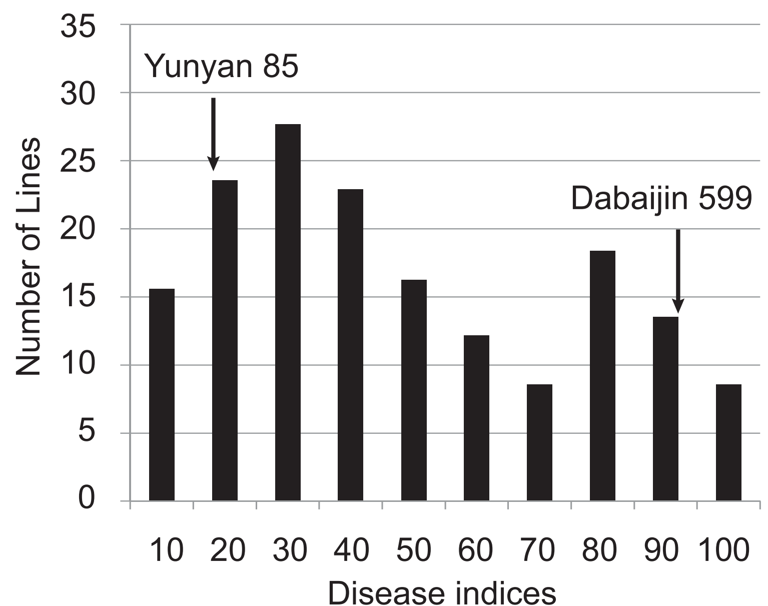

Fig. 2 Frequency distribution for the RIL population exhibiting disease indices in greenhouse environment. The disease indices for parental lines are indicated by arrows

Table 1 ANOVA analysis of variance for disease indices for RIL population in five environments

Source	*df*	Mean squares	*F*-value	*P* value
Environment	4	68,378.49	1240.05	< 0.0001
Genotype (lines)	178	5038.72	91.37	< 0.0001
Environment × block	6	3461.44	62.77	< 0.0001
Environment × genotype	712	535.46	9.71	< 0.0001
Genotype × block	534	59.50	1.07	0.152

df, degrees of freedom

In total, 2,404,474 SNPs were detected using GATK software. For the SNP loci, two to four alleles were identified as polymorphic and considered potential markers. Across the whole RIL population, 460,210 polymorphic markers were successfully identified. The SNP markers were classified into four genotypes: aa × bb, hk × hk, lm × ll, and nn × np, where aa × bb indicates that both the parents were homozygous, with the genotype of one parent being aa and the other being bb; lm × ll and nn × np indicate that one parent was homozygous and the other was heterozygosis, and the hk × hk indicates that both the parents were heterozygosis. The numbers of SNPs classified into these four marker types were 150,489; 120,264; 87,089; and 102,368, respectively. Only 150,489 SNPs with parental homozygous genotype (aa × bb) were used for further analyses. The markers with low sequence depths, missing in the parents or RIL population, or having segregation distortion, were filtered. The remaining 7878 markers were used to construct the genetic map (Additional file 1: Table S1).

Construction of a high-density genetic map

In total, 7734 SNP markers were mapped to 24 LGs (Fig. 4, Additional file 2: Table S2). The number of LGs corresponded to the number of chromosomes in tobacco ($x = 24$). The genetic map contained 2800 recombination loci with 63.79% co-segregation markers. The map spanned a total distance of 2689.06 cM, with an average marker interval of 0.35 cM (Table 2). LG3 was the largest group, containing 1230 markers that spanned 190.83 cM, with an average marker interval of 0.16 cM. The shortest linkage group was LG4, which harbored 184 markers and had a genetic length of 53.82 cM with an average inter-marker distance of 0.29 cM. The degree of linkage between markers was reflected by "Gaps ≤ 5" indicating the percentage of gaps in which the distance between adjacent markers was smaller than 5 cM. The value ranged from 95.28 to 100.00%, with an average value of 98.61%. The largest gap on this map was 16.74 cM located on LG12. Tobacco is an allotetraploid ($2n = 4x = 48$) and is most likely the result of a tetraploidization inter-specific hybrid between *N. sylvestris* (S-genome) and *N. tomentosiformis* (T-genome). According to the alignment with the two diploid ancestral genomes of common tobacco (Sierro et al. 2013), 61.09% markers were mapped to the S-genome and 35.75% to the T-genome. Several LGs contained both S- and T-genome-specific markers. The haplotype and heat maps reflect the map quality (Additional files 3 and 4: Figures S1 and S2, respectively).

QTL analysis

Ten QTLs with additive effects were significantly associated with the DI of black shank of tobacco,

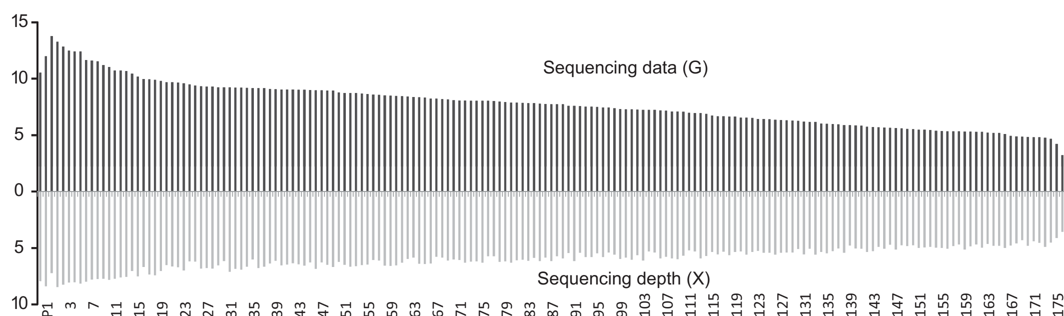

Fig. 3 Sequencing data and depth statistics of parental lines and each individual of the RIL population. The *x*-axis represents the parental lines and each individual of the RIL population, and the *y*-axis represents the total numbers of bases (Gb; black) and average sequencing depth (X; gray)

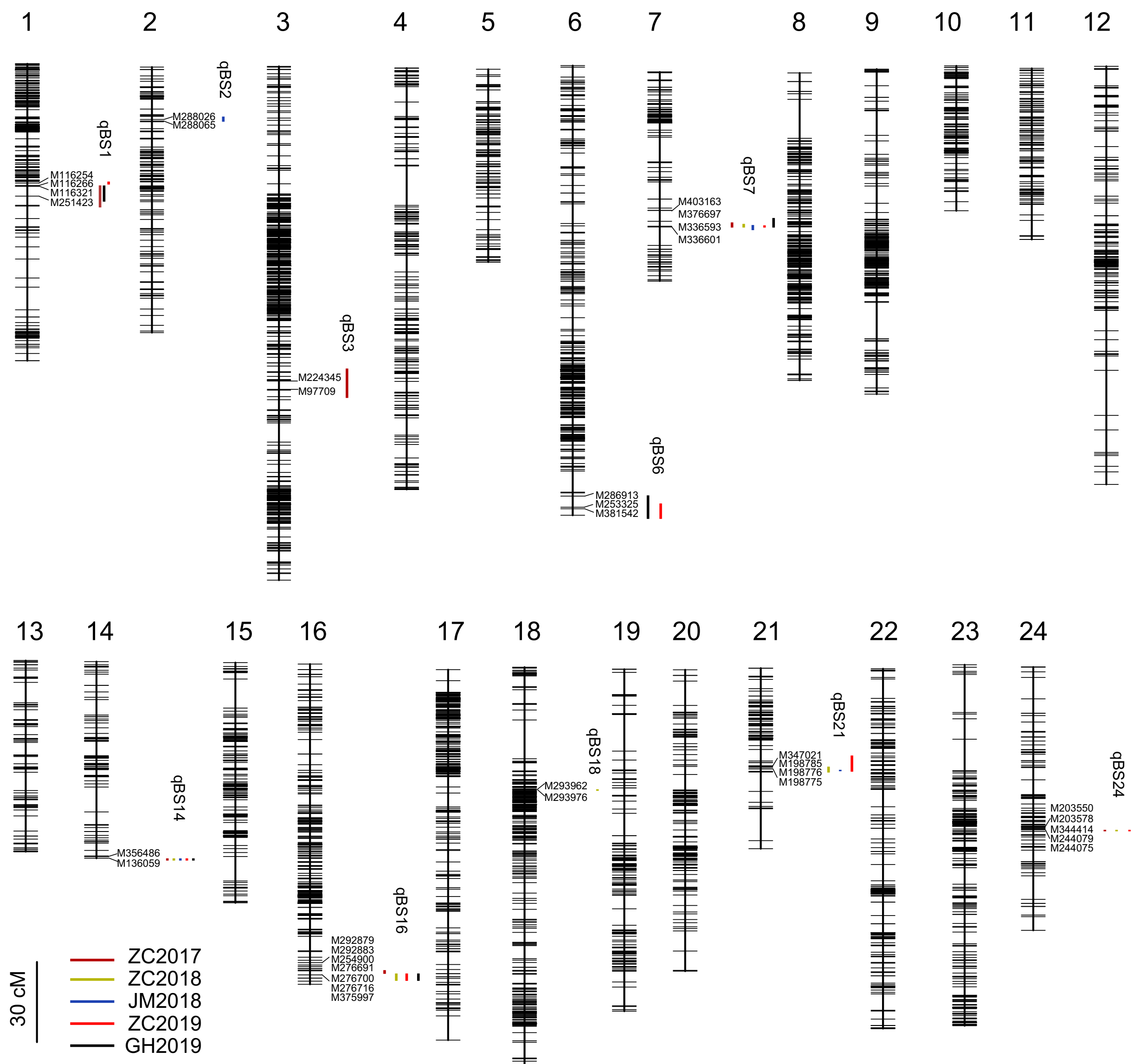

Fig. 4 High-density genetic map of tobacco constructed by SNP markers. The markers position and name in Table S2. Marker names of the QTL identified on each chromosome are given on the right. The vertical colored bars represent QTLs for four field and greenhouse under the disease progress. Length of vertical colored bars indicates confidence interval calculated with a support interval of 1 LOD score

having LOD scores > 2.0 (Table 3) over five different environments. These QTLs were mapped to 10 link groups. Among those QTL, two stable QTL named *qBS7* and *qBS14* were repeatedly detected to resistance black shank under all five environments. The QTL *qBS7* located in LG7 was a major QTL and explained the greatest observed phenotypic variations ranging from 16.48 to 62.20%. High LOD score values of 28.33, 13.75, 43.40, 7.54, and 38.72, respectively, were observed for *qBS7* in the five environments. QTL *qBS7* in ZC2017, ZC2018, and ZC2019 was flanked by markers M376697 and M336593 at ~1.5 cM, by markers M403163 and M376697 at 4 cM in GH2019, and by M336593 and M336601 at 2 cM in JM2018. QTL *qBS14* was identified in LG14, explaining 3.94-11.29% of the phenotypic variation in the five different environments. QTL *qBS14* was flanked by two markers,

Table 2 Distribution of mapped markers on the 24 linkage groups in tobacco

Chromosome number	Marker number	Total distance (cM)	Average distance (cM)	Largest gap (cM)	% of gap (< 5 cM)
1	406	110.28	0.27	8.42	99.51
2	184	98.63	0.54	4.02	100
3	1230	190.83	0.16	7.12	99.84
4	184	53.82	0.29	3.78	100
5	138	71.6	0.52	4.33	100
6	528	167.02	0.32	7.83	99.43
7	181	77.62	0.43	9.22	98.33
8	525	114.18	0.22	14.47	99.62
9	275	120.71	0.44	13.58	98.54
10	315	156.49	0.5	14.64	98.09
11	119	63.55	0.53	5.62	99.15
12	218	155.31	0.71	16.74	96.31
13	88	70.96	0.81	7.57	96.55
14	94	73.05	0.78	9.23	96.77
15	263	89.21	0.34	10.78	99.24
16	357	118.94	0.33	5.4	99.44
17	799	137.58	0.17	8.98	99.62
18	437	147.78	0.34	14.41	99.31
19	192	127.01	0.66	13.64	96.86
20	393	111.93	0.28	9.05	98.98
21	170	67.07	0.39	11.84	98.82
22	269	133.65	0.5	7.19	98.51
23	262	134.05	0.51	14.14	98.47
24	107	97.79	0.91	8.76	95.28
Total	7734	2689.06	0.35	16.74	98.61

M356486 and M136059, at 0.75 cM. In addition, the *qBS16* was a minor QTL that explained 2.95-5.31% of phenotypic variation in four environments, with 2.70-3.39 LOD score of values. Minor QTL *qBS1*, *qBS21*, and *qBS24* can be found in three environments. QTL qBS1 was detected not only in the field including ZC2017 and ZC2019 but also in the greenhouse in GH2019, explaining 2.46%, 2.33%, and 3.33% of the phenotypic variance while *qBS21* and *qBS24* were just found in field surroundings. In additional, minor QTL *qBS6* with lower PVE and LOD value were found in environments of GH2019 and ZC2019. The remaining three minor QTL *qBS2*, *qBS3*, and *qBS18* were just detected in one environment. 'Yunyan 85' contributed the favorable allele at each of the five genomic regions while the favorable alleles at five of the 10 QTL positions were from the Dabaijin 599 parent.

Identification of the candidate genes

Four markers (marker376697, marker336593, marker336601, and marker403163) and two markers (marker356486 and marker136059) were located in QTLs *qBS7* and *qBS14*, respectively. Unfortunately, these six markers were not mapped to the physical map of the tobacco genome but were merely located on five unanchored scaffolds. By overlap extension, combined with four genome sequences ('TN90', 'K326', *N. sylvestris*, and *N. tomentosiformis*) (Edwards et al. 2017; Sierro et al. 2013; Sierro et al. 2014), we constructed 3.9 MB and 2.8 MB of scaffold sequences near the five markers for *qBS7* and *qBS14*, respectively (Table S3). According to the gene annotation information for 'K326', 72 and 66 candidate genes were found within the confidence intervals of *qBS7* and *qBS14* respectively (Additional file 5: Table S3). Among these 138 predicted genes,

fourteen candidate genes encoded nucleotide-binding site-leucine-rich repeat (NBS-LRR) proteins associated with plant defense responses to pathogens. Each of the two QTLs has one candidate gene that encodes a protein with high homology to downy mildew disease resistance protein recognition of *Peronospora parasitica* 13 (RPP13) in *Arabidopsis thaliana* (Bittner-Eddy et al. 2000). In total, eleven candidate genes encode proteins with high homology to the late blight resistance protein R1 in potato (Ballvora et al. 2002). Five candidate genes encode glycine-rich protein (GRP) and two candidate genes encode cytochrome P450 protein (CYP). Besides, Nitab4.5_0002781g0070.1 encodes serine/threonine-protein kinase (STK) protein with LRR receptor.

Discussion

A high-density linkage map is an important tool for genomic analyses, QTL mapping, and marker-assisted breeding. The first high-density linkage map in tobacco was constructed using 2318 microsatellite markers and covered 3270 cM (Bindler et al. 2011) based on an F_2 mapping population derived from the intervarietal cross of 'Hicks Broadleaf' × 'Red Russian'. However, few SSR markers can be used for QTL mapping for resistance to black shank in the following population. In total, 156, 373, 204, and 258 SSR markers were used in 'Beinhart 1000-1' × 'Xiaohuangjin 1025' (Zhang et al. 2018), 'Florida 301' × 'Hicks' (Xiao et al. 2013), 'Beinhart 1000' × 'Hicks' (Vontimitta and Lewis 2012a; Vontimitta and Lewis 2012b), and 'K346' × 'TI 1068' (Drake-Stowe et al. 2017), respectively. These maps have low density levels and even the gaps divide the chromosomes into two or more LGs. The insufficient SSR marker numbers do not allow the construction of SSR marker-based maps to finely map QTLs.

With the development of high-throughput sequencing technologies, SNP markers have been applied to genetic map construction, improving the coverage and resolution of genetic maps even from populations of flue-cured tobacco crosses. Gong et al. (2016) used SNP markers and SLAF-seq to construct a map that harbored 4215 markers, with a total distance of 2662.43 cM and an average inter-marker distance of 0.60 cM (Gong et al. 2016). Here, we constructed a high-density genetic map using the SNP markers developed through a RAD-seq method. In total, 7734 SNPs were assigned to 24 LGs corresponding to the 24 chromosome pairs of the tobacco genome. The map contained a greater number of markers and was more saturated than the previous two maps, with an average inter-marker distance of 0.35 cM. The total covered distance was approximately the same as that of the SLAF-based SNP map and slightly less than that of the SSR map (Bindler et al. 2011). Nevertheless, the saturated SNP genetic map had 63.79% co-segregated markers and represented 2800 recombinant loci. Several gaps > 10 cM can be seen on the genetic map (Fig. 4). Because of a large number of repetitive or homologous sequences in allotetraploid tobacco and low sequencing depths, many true SNPs are being eliminated, which might cause the gaps.

The inheritance of the high level of partial resistance to *P. nicotianae* in 'Florida 301' and 'Beinhart 1000' is complex and controlled by polygene (Vontimitta and Lewis 2012b; Xiao et al. 2013). The current QTL study suggests that the resistance of 'Yunyan 85' is controlled by one major and several minor gene loci. The favorable alleles contributed by 'Yunyan 85' were at five of the 10 QTL. Meanwhile, the black shank-susceptible cultivar 'Dabaijing 599' may contribute five minor resistance QTLs. 'Yunyan 85' × 'Dabaijin 599' RIL exhibited significantly greater resistance of black shank than 'Yunyan 85'.

Using the high-density SNP genetic map, the accuracy of QTL mapping for resistance to black shank in tobacco improved. Our fine mapping results in five distinct environments revealed that three stable QTLs, *qBS7*, *qBS14*, and *qBS16*, were within a tight interval. QTL *qBS7*, a stable major QTL, had the greatest effects and controlled the largest percentage of the phenotypic variation for black shank resistance (16.48-62.20%) in 'Yunyan 85'. This QTL also had the greatest effect and controlled the largest percentages of the phenotypic variation for black shank resistance in 'Florida 301' (16.9%) (Xiao et al. 2013), 'K346' (38.0%) (Drake-Stowe et al. 2017), and 'Beinhart 1000' (25.4%) (Vontimitta and Lewis 2012a; Vontimitta and Lewis 2012b). In the present study, *qBS7* was confirmed and the confidence interval was narrowed to 1-4 cM, which is a significant improvement compared with previous reports (Vontimitta and Lewis 2012a; Vontimitta and Lewis 2012b; Xiao et al. 2013; Drake-Stowe et al. 2017). Recently, the QTL was localized to within a genetic interval of approximately 3.51 cM between marker Nt1AF8640 and marker Nt1AD1494 using sub-NILs (Ma et al. 2019). In our result, the interval of

Table 3 QTLs of tobacco black shank resistance detected in the mapping population

QTL	Environment	Chromosome	Position	Left marker	Right marker	LOD	PVE (%)	Add	Left CI	Right CI
qBS7	ZC2017	7	56.5	M376697	M336593	28.33	46.85	22.68	55.25	57.25
	GH2019	7	55	M403163	M376697	13.75	26.62	14.26	53.25	57.25
	ZC2018	7	56.5	M376697	M336593	43.40	62.2	31.63	55.75	57.25
	JM2018	7	57.0	M376697	M336601	7.54	16.48	8.32	56.25	59.25
	ZC2019	7	56.5	M376697	M336593	38.72	55.13	19.39	55.75	56.75
qBS14	ZC2017	14	73	M356486	M136059	3.38	3.94	− 6.46	72.25	73
	GH2019	14	73	M356486	M136059	6.26	11.49	− 11.06	72.25	73
	ZC2018	14	73	M356486	M136059	4.68	5.11	− 7.24	72.25	73
	JM2018	14	72.5	M356486	M136059	3.23	6.67	− 5.23	72.25	73
	ZC2019	14	73	M356486	M136059	7.53	6.91	− 6.74	72.25	73
qBS16	ZC2017	16	111.5	M292879	M276716	2.70	3.58	− 6.29	110.75	113.75
	GH2019	16	115	M276716	M375997	2.90	5.31	− 6.65	113.75	116.75
	ZC2018	16	115	M276716	M375997	2.73	4.52	− 5.78	113.75	116.75
	ZC2019	16	115.5	M375997	M336467	3.39	2.95	− 4.39	113.75	116.75
qBS24	ZC2017	24	59	M203550	M203578	2.98	3.47	6.04	58.75	59.75
	ZC2018	24	60	M244079	M244075	3.61	3.09	6.72	59.75	60.25
	ZC2019	24	60	M244079	M244075	2.82	2.56	4.09	59.75	60.25
qBS1	ZC2017	1	47.5	M116321	M251423	2.04	2.46	− 5.1	45.25	54.25
	GH2019	1	47.5	M116321	M251423	2.88	3.33	− 5.98	45.25	51.75
	ZC2019	1	44.5	M116254	M116266	2.50	2.23	− 2.85	43.75	45.75
qBS21	ZC2018	21	37	M198776	M198775	2.52	2.52	− 3.95	35.25	37.75
	JM2018	21	37	M198776	M198775	2.72	3.34	− 3.68	36.75	37.25
	ZC2019	21	36.5	M347021	M198785	2.58	2.28	− 2.89	31.25	37.75
qBS6	GH2019	6	163.5	M286913	M253325	2.39	4.34	− 5.93	158.75	167
	ZC2019	6	164.5	M253325	M381542	2.73	2.42	− 3.04	160.75	167
qBS2	JM2018	2	19.5	M288026	M288065	2.94	6.04	4.95	18.25	20.25
qBS18	ZC2018	18	47.5	M293962	M293976	2.19	2.17	4.72	47.25	47.75
qBS3	ZC2017	3	117.5	M224345	M97709	2.17	2.76	5.51	111.75	122.25

LOD maximum log-likelihood score (QTL peak), *PVE* percentage of the variance explained by the QTL, *ADD* additive effects

M336601-M403163 was the same as Nt1AF8640-Nt1AD1494 based upon mapping the marker sequences with physical sequence of K326 reference genome. In the interval, our M376697 and four reported markers Nt1AF0086, Nt1AG1690, NtAB1658, and Nt1AA1707 were most highly associated with resistance in nearby position. These markers can be useful for marker-assisted selection in breeding programs at the *qBS7* locus. In addition, the QTL was identified as being associated with resistance to bacterial wilt (Drake-Stowe et al. 2017). 'Yunyan 85' exhibited a low-to-moderate level of resistance to bacterial wilt in the field. It is valuable for fine mapping and cloning the gene(s) underlying the *qBS7* QTL. The second favorable allele contributed by 'Yunyan 85' was in QTL *qBS1* in three environments with smaller effect. The other three minor QTLs with positive additive effects were only identified in one environment.

The second-effect QTL *qBS14* was also found to have the same location as the third QTL identified from 'Florida 301' and 'K346' based on SSR primer sequences and SNP sequence mapping to the 'K326' physical map (Xiao et al. 2013; Drake-Stowe et al. 2017). The QTL explained 6.82% of the average phenotypic variation of black shank resistance in the five environmental conditions in 'Yunyan 85' × 'Dabaijin 599' with 0.75-cM confidence interval, while it controlled 6.0% and 4.4% of the variation in a 'Florida

301' × 'Hicks' and 'K346' × 'TI1068' RIL mapping populations with more than 6-cM confidence interval, respectively. The QTL had a minor effect on black shank resistance and no association with resistance to bacterial wilt. The favorable allele was contributed by the black shank-susceptible parent 'Dabaijing 599'. The pedigree information suggests that the flue-cured variety 'Yunyan 85' possesses polygenic black shank resistance derived from cigar tobacco 'Florida 301'. It is possible that 'Yunyan 85' lost many minor QTLs from 'Florida 301', such as *qBS14*. The other four QTLs *qBS*1, *qBS*6, *qBS*21, and *qBS*24 with negative additive effects have lower LOD values and PVEs. Nevertheless, these QTLs were identified in multiple environments. Markers associated with these QTLs could be useful for transferring this favorable allele to increase the resistance of 'Yunyan 85' or other varieties.

In total, 138 candidate genes for black shank resistance were predicted in the target area. Of these, 14 were typical NBS-LRR resistance (*R*) genes and several were defense-related genes. *R* genes confer very high levels of narrow spectrum race-specific resistance to plant disease, but major *R* genes are lack of durability in some systems (McDonald and Linde 2002). Few *R* genes to date confer partial pathogen resistance. However, Jiang et al. reveal for the first time that NB-LRR gene *R8* conferred broad spectrum and durable field resistance against *P. infestans* (Jiang et al. 2018). Apart from the typical NBS-LRR resistance protein, others such as serine/threonine-protein kinase (STK), glycine-rich protein (GRP), and cytochrome P450 protein (CYP) also involve in defense-responsive or defense-related against pathogen attack in crops. Further fine mapping will be required to reduce the number of candidate genes. This will be necessary to develop additional markers and to use larger populations to construct a fine physical map. The candidate gene expression data based on RNA-seq could also be useful for identifying differentially expressed gene. Moreover, the CRISPR/Cas9 gene editing technology can be used to inactivate candidate genes to validate functions.

Conclusions

This research reported a high-density SNP genetic linkage map of flue-cured tobacco using a RIL population and SNP markers developed by RAD-seq. The genetic map had 7734 markers and 2689 cM, with an average marker interval of 0.35 cM. We also identified 10 QTLs associated with resistance to black shank across five environments. Furthermore, 138 candidate genes of two major QTLs were identified in the confidence intervals. The result of this research will be helpful in elucidating the mechanism of disease resistance and in marker-assisted selection.

Author contributions YS, YZ, and XX collected the plant materials used in this study. DG and MC carried out the determination of resistance to black shank. DG, MC, and XZ analyzed data and prepared the manuscript. DG and XX planned and supervised this work and edited the manuscript. All authors read and approved the final manuscript.

Funding information This work was supported by the Fundamental Research Funds for Central Non-profit Scientific Institution (1610232017009 and 1610232017012) and the Science Foundation for Young Scholars of Tobacco Research Institute of Chinese Academy of Agricultural Sciences (2015B01).

Compliance with ethical standards

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

References

Apple JL (1962) Transfer of resistance to black shank (*Phytophthora parasitica* var *nicotianae*) from *Nicotiana plumbaginifolia* to *N. tabacum*. Phytopathology 52:1

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3(10):e3376. https://doi.org/10.1371/journal.pone.0003376

Ballvora A, Ercolano MR, Weiss J, Meksem K, Bormann CA, Oberhagemann P, Salamini F, Gebhardt C (2002) The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J 30(3):361-371

Bindler G, Plieske J, Bakaher N, Gunduz I, Ivanov N, Van der Hoeven R, Ganal M, Donini P (2011) A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor Appl Genet 123(2):219-230. https://doi.org/10.1007/s00122-011-1578-8

Bittner-Eddy PD, Crute IR, Holub EB, Beynon JL (2000) RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in Peronospora parasitica. Plant J 21(2):177-188. https://doi.org/10.1046/j.1365-313x.2000.00664.x

Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA, Martin GB (2012) A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. Mol Plant Microbe Interact 25(12):1523-1530. https://doi.org/10.1094/MPMI-06-12-0148-TA

Chaplin JF (1962) Transfer of black shank resistance from *Nicotiana plumbaginifolia* to flue-cured *N. tabacum*. Tob Sci 6:184-189

Chaplin JF (1966) Comparison of tobacco black shank (*Phytophthora parasitica* var. *nicotianae*) resistance from four sources. Tob Sci 10:55-58

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43(5): 491-498. https://doi.org/10.1038/ng.806

Drake K, Lewis RS (2013) An introgressed *Nicotiana rustica* genomic region confers resistance to *Phytophthora nicotianae* in cultivated tobacco. Crop Sci 53(4):1366-1374. https://doi.org/10.2135/cropsci2012.10.0605

Drake KE, Moore JM, Bertrand P, Fortnum B, Peterson P, Lewis RS (2015) Black shank resistance and agronomic performance of flue-cured tobacco lines and hybrids carrying the introgressed *Nicotiana rustica* region, Wz. Crop Sci 55(1): 79-86. https://doi.org/10.2135/cropsci2014.02.0164

Drake-Stowe K, Bakaher N, Goepfert S, Philippon B, Mark R, Peterson P, Lewis RS (2017) Multiple disease resistance loci affect soilborne disease resistance in tobacco (*Nicotiana tabacum*). Phytopathology 107(9):1055-1061

Edwards KD, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans AD, Bombarely A, Allen F, Hurst R, White B, Kernodle SP, Bromley JR, Sanchez-Tamburrino JP, Lewis RS, Mueller LA (2017) A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. BMC Genomics 18(1):448. https://doi.org/10.1186/s12864-017-3791-6

Gallup CA, Shew HD (2010) Occurrence of race 3 of *Phytophthora nicotianae* in North Carolina, the causal agent of black shank of tobacco. Plant Dis 94:557-562

Gong D, Huang L, Xu X, Wang C, Ren M, Wang C, Chen M (2016) Construction of a high-density SNP genetic map in flue-cured tobacco based on SLAF-seq. Mol Breed 36(7):100

Heggestad H, Lautz W (1957) Some results of studies on resistance to tobacco black shank. Phytopathology 47:452-452

Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D, Lu Y, Weng Q, Liu K, Zhou T, Jing Y, Si L, Dong G, Huang T, Lu T, Feng Q, Qian Q, Li J, Han B (2011) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44(1):32-39. https://doi.org/10.1038/ng.1018

Jansen J, Jong A, Ooijen J (2001) Constructing dense genetic linkage maps. Theor Appl Genet 102:10

Jiang R, Li JC, Tian ZD, Du J, Armstrong M, Baker K, Lim JTY, Vossen JH, He H, Portal L, Zhou J, Bonierbale M, Hein I, Lindqvist-Kreuze H, Xie CH (2018) Potato late blight field resistance from QTL dPI09c is conferred by the NB-LRR gene R8. J Exp Bot 69(7):1545-1555. https://doi.org/10.1093/jxb/ery021

Kosambi D (1944) The estimation of map distance from recombination values. Ann Eugenics 12:172-175

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754-1760. https://doi.org/10.1093/bioinformatics/btp324

Li H, Ye G, Wang J (2007) A modified algorithm for the improvement of composite interval mapping. Genetics 175(1):361-374. https://doi.org/10.1534/genetics.106.066811

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Liu D, Ma C, Hong W, Huang L, Liu M, Liu H, Zeng H, Deng D, Xin H, Song J, Xu C, Sun X, Hou X, Wang X, Zheng H (2014) Construction and analysis of high-density linkage map using high-throughput sequencing data. PLoS One 9(6):e98855. https://doi.org/10.1371/journal.pone.0098855

Ma JM, Heim C, Humphry M, Nifong JM, Lewis RS (2019) Genetic analysis of Phn7.1, a major QTL conferring partial resistance to Phytophthora nicotianae in Nicotiana tabacum. Mol Breed 39(1). https://doi.org/10.1007/s11032-018-0923-x

McCorkle KL, Drake-Stowe K, Lewis RS, Shew D (2018) Characterization of *Phytophthora nicotianae* resistance conferred by the Introgressed *Nicotiana rustica* region, Wz, in flue-cured tobacco. Plant Dis 102(2):309-317. https://doi.org/10.1094/Pdis-03-17-0339-Re

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. Annu Rev Phytopathol 40:349. https://doi.org/10.1146/annurev.phyto.40.120501.101443

McIntyre JL, Taylor GS (1978) Race 3 of *Phytophthora parasitica* var. *nicotianae*. Phytopathology 68:35-38

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

Nielsen M (1992) Sources of resistance to black shank and black root rot diseases. CORESTA Inf Bull 3(4):144-150

Sierro N, Battey JN, Ouadi S, Bovet L, Goepfert S, Bakaher N, Peitsch MC, Ivanov NV (2013) Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Genome Biol 14(6):R60. https://doi.org/10.1186/gb-2013-14-6-r60

Sierro N, Battey JND, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch MC, Ivanov NV (2014) The tobacco genome sequence and its comparison with those of tomato and potato. Nat Commun 5(5):3833. https://doi.org/10.1038/Ncomms4833

Sullivan MJ, Melton TA, Shew HD (2005) Managing the race structure of *Phytophthora parasitica* var. *nicotianae* with cultivar rotation. Plant Dis 89:1285-1294

Tan C, Li Y, Wang Y, Ma W, Lei Y (1997) Breeding and selecting a new variety of flue-cured tobacco Yunyan 85 and its characteristics. Chin Tobacco Sci 1:7-10

Tisdale WB (1931) Development of strains of cigar wrapper tobacco resistant to black shank (*Phytophthora nicotianae* Breda de Haan). Florida Univ Agr Expt Stat Bull 226:1-45

Valleau WD, Stokes GW, Johnson EM (1960) Nine years experience with the *Nicotiana longiflora* factor for resistance to

Phytophthora parasitica var. *nicotianae* in the control of black shank. Tob Sci 4:92-94

van Jaarsveld E, Wingfield MJ, Drenth A (2002) Evaluation of tobacco varieties for resistance to races of *Phytophthora nicotianae* in South Africa. J Phytopathol 150:456-462

Van Ooijen JW (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet Res 93(5):343-349. https://doi.org/10.1017/S0016672311000279

van Os H, Stam P, Visser RG, van Eck HJ (2005) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theor Appl Genet 112(1):187-194. https://doi.org/10.1007/s00122-005-0124-y

Vontimitta V, Lewis RS (2012a) Growth chamber evaluation of a tobacco 'Beinhart 1000' × 'Hicks' mapping population for quantitative trait loci affecting resistance to multiple races of *Phytophthora nicotianae*. Crop Sci 52(1):91-98

Vontimitta V, Lewis RS (2012b) Mapping of quantitative trait loci affecting resistance to *Phytophthora nicotianae* in tobacco (*Nicotiana tabacum L.*) line Beinhart-1000. Mol Breed 29(1):89-98

West MA, van Leeuwen H, Kozik A, Kliebenstein DJ, Doerge RW, St Clair DA, Michelmore RW (2006) High-density haplotyping with microarray-based expression and single feature polymorphism markers in Arabidopsis. Genome Res 16(6):787-795. https://doi.org/10.1101/gr.5011206

Wills W (1971) A comparison of black shank resistance of tobacco measured in inoculated whole plants and excised leaf tissue. Tob Sci 15:47-50

Xiao BG, Drake K, Vontimitta V, Tong ZJ, Zhang XT, Li MY, Leng XD, Li YP, Lewis RS (2013) Location of genomic regions contributing to *phytophthora nicotianae* resistance in tobacco cultivar Florida 301. Crop Sci 53(2):473-481. https://doi.org/10.2135/cropsci2012.06.0376

Zhang YS, Guo X, Yan XX, Ren M, Jiang CH, Cheng YZ, Wen LY, Liu D, Zhang Y, Sun MM, Feng QF, Yang AG, Cheng LR (2018) Identification of stably expressed QTL for resistance to black shank disease in tobacco (*Nicotiana tabacum* L.) line Beinhart 1000-1. Crop J 6(3):282-290. https://doi.org/10.1016/j.cj.2017.12.002

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.