Australian Journal of Crop Science AJCS

AJCS 10(8):1188-1196 (2016) ISSN:1835-2707
DOI: 10.21475/ajcs.2016.10.08.p7893

SNP association analysis of resistance to Verticillium wilt (*Verticillium dahliae* Kleb.) in spinach

Ainong Shi[1*], Beiquan Mou[2*], Jim Correll[3], Dennis Motes[4], Yuejin Weng[1], Jun Qin[1], Wei Yang[1]

[1]Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA
[2]Crop Improvement and Protection Research Unit, US Department of Agriculture, Agricultural Research Service (USDA-ARS), 1636 E. Alisal Street, Salinas, CA 93905, USA
[3]Department of Plant Pathology, University of Arkansas, Fayetteville, AR 72701, USA
[4]Vegetable Research Center, University of Arkansas, Alma, AR 72921, USA

***Corresponding author: ashi@uark.edu; Beiquan.Mou@ARS.USDA.GOV**

Abstract

Verticillium wilt, caused by *Verticillium dahliae* Kleb., is an important disease of spinach (*Spinacia oleracea* L.) and use of genetic resistance is the most economical method of controlling this disease. The objective of this research was to conduct molecular association analysis for Verticillium wilt resistance in spinach. A total of 95 USDA spinach accessions were evaluated for resistance to Verticillium wilt in this study. Phenotyping was conducted using a 0-4 scale of disease severity scores of Verticillium wilt and genotyping was performed using 2,878 SNPs which were postulated from genotyping by sequencing (GBS). STRUCTURE 2.3.4 and MEGA 6 were used for population structure and genetic diversity analysis. The single marker regression (SMR) from QGene, general linear mode (GLM) and mixed linear mode (MLM) from TEASSEL, and compressed mixed linear model (cMLM) and enriched compressed mixed linear model (EcMLM) from GAPIT were used for association analysis of Verticillium wilt resistance. Significant genetic variation of Verticillium wilt disease resistance was observed among the 95 spinach accessions with a wide range from 0.3 to 3.0 on a 0-4 scale. Two well-differentiated genetic populations and admixtures were postulated in the spinach panel. Five SNP markers, AYZV02052595_108, AYZV02112284_14543, AYZV02123399_146, AYZV02164612_331, and AYZV02170942_274 were identified to be associated with Verticillium wilt resistance with R-squared values from 9.3 to 18.2%. These markers may provide a tool utilized in molecular spinach breeding to select Verticillium wilt resistance through marker-assisted selection.

Keywords: spinach, *Spinacia oleracea*, Verticillium wilt, *Verticillium dahliae*, association mapping, single nucleotide polymorphism, genotyping by sequencing.
Abbreviations: SNP_single nucleotide polymorphism, GBS_genotyping by sequencing, SMR_single marker regression, GLM_general linear model, MLM_mixed linear mode, cMLM_compressed mixed linear model, EcMLM_enriched compressed mixed linear model, MAS_marker-assisted selection.

Introduction

Spinach (*Spinacia oleracea* L.) is an economically important vegetable crop worldwide (Morelock and Correll, 2008; Correll et al., 2011). In addition to its economic importance, spinach is one of the faster growing vegetable crops in the US and other regions in terms of per capita consumption and is considered one of the healthiest vegetables in the human diet due to its high concentration of nutrients and health-promoting compounds (Dicoteau, 2000; Morelock and Correll, 2008). Diseases represent a significant constraint in spinach production (Correll et al., 1994). Verticillium wilt, caused by *Verticillium dahliae* Kleb., is one of the most important fungus diseases in spinach seed production areas (du Toit et al., 2005; Maruthachalam et al., 2013; Mou et al., 2015; Villarroel-Zeballos et al., 2012). *V. dahliae* is a soilborne pathogen that has a wide host range causing vascular wilt on a number of economically important crops including bell pepper, cabbage, cauliflower, chili pepper, cotton, eggplant, lettuce, mint, potato, strawberry, and tomato (Bhat and Subbarao, 1999; Maruthachalam et al., 2013; Villarroel-Zeballos et al., 2012). The pathogen can be transmitted by vegetative material, farm equipment, water, soil, and air and is also a seedborne pathogen in spinach, lettuce and other crops (Correll et al., 1994; du Toit et al., 2005; Maruthachalam et al., 2013; Vallad et al., 2006; van der Spek, 1972). Maruthachalam et al. (2013) indicated that seedborne transmission of *V. dahliae* in spinach is a serious concern, especially if the transmitted isolates become established in coastal California, where lettuce and other hosts are rotated with spinach, and where approximately 69.3% of the total fresh market spinach in the United States was produced in 2014 (http://usda.mannlib.cornell.edu/usda/current/VegeSumm/VegeSumm-01-29-2015.pdf). *V. dahliae* was recorded to infect spinach in 1982 (Sackston and Sedun), reported to infect spinach and lettuce field in the Salinas Valley area in 1995 (Bryant, 2009), however it was realized as an important disease of spinach in 2005 (du Toit et al., 2005; Iglesias-Garcia et al., 2013; Villarroel-Zeballos et al., 2012). More

recent surveys of commercial spinach seed lots indicated that up to 85% of the seeds per lot are infested with *V. dahliae* (Duressa et al., 2012). Although Verticillium wilt has not been a problem for spinach production in California due to the late appearance of symptoms after the stem elongation (bolting) stage, infected spinach seeds introduce or increase inoculum in soils for rotational crops such as lettuce (Short et al., 2015).

A wide range of management options are available to help control Verticillium wilt in spinach such as rotation and seed treatment by chemicals for decreasing the disease (du Toit et al., 2005; Maruthachalam et al., 2013; Vallad et al., 2006; Villarroel-Zeballos et al., 2012). But because *V. dahliae* is a soil-borne pathogen and can survive in soil for years, it is very difficult to manage through rotation or chemicals (Maruthachalam et al., 2013; Villarroel-Zeballos et al., 2012). However, the use of host resistance will be the most effective and economical practice to control *V. dahliae* (Mou et al., 2015; Maruthachalam et al., 2013). Villarroel-Zeballos et al. (2012) evaluated 120 USDA spinach germplasm lines and 10 commercial spinach hybrids for resistance to *V. dahliae* under greenhouse conditions and a wide range in disease severity was observed. Many accessions were susceptible, but some quantitative resistance was observed among the spinach accessions. Four spinach accessions were consistently rated as partially resistant to Verticillium wilt (disease severity 25% or less): PI 175931, Ames 26243, PI 163309, and PI 261789 (Villarroel-Zeballos et al., 2012). Beiquan et al. (2015) screened 268 USDA spinach germplasm accessions plus 9 commercial cultivars; reported 18 accessions plus two cultivars for resistance to two isolates So 923 and So 925 of *V. dahliae* race 2 and one isolate So 302 of race 1; and observed that there was a wide range in disease incidence and severity of resistance to the three isolates.

So far, although the genetic basis has not been characterized, Villarroel-Zeballos et al. (2012) and Mou et al. (2015) did describe some levels of resistance. Villarroel-Zeballos et al. (2012) evaluated 130 spinach genotypes and recognized that the Verticillium wilt resistance in spinach is a quantitative trait. It would be time-consuming to transfer these quantitative traits through classic plant breeding approaches. However, molecular plant breeding can be an efficient way to select quantitative traits through marker assisted selection (MAS). Single nucleotide polymorphism (SNP), with its abundance, cost efficiency and high-throughput scoring, has become a powerful tool in genome mapping, association studies, diversity analysis, and tagging of important genes in plant genomics (Collard and Mackill, 2008; Xu and Crouch, 2008). Therefore, identification of SNP markers associated with Verticillium wilt resistance will provide breeders with a powerful tool to assist in selecting for disease resistance in spinach breeding programs. Genotyping by sequencing (GBS) is one of the next-generation sequencing platforms to discover SNPs without prior knowledge of the genome in spinach (Elshire et al., 2011; He et al., 2014; Poland and Rife, 2012; Sonah et al., 2013). The spinach genome sequences AYZV01 and AYZV02 are available to the public (http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV01 and http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV02), and represent approximately half of the spinach genome (Dohm et al., 2014; Minoche et al., 2015). In addition, a more comprehensive version of the spinach genome assembly will be made publicly available in 2016 (van Deynze, 2014; van Deynze et al., 2015; Allen van Deynze, personal communication). These resources provide a reference for SNP discovery and association analysis in spinach. The objective of this research was to conduct association analysis and identify SNP markers associated with Verticillium wilt resistance in the USDA spinach collection.

Results

Phenotyping of Verticillium wilt resistance

The Verticillium disease severity varied among the 95 spinach genotypes tested, ranging from 0 to 3, and averaged 2.1; the standard deviation (stdv) was 0.63 with stdev error 0.007, indicating that there were significant genetic differences of Verticillium disease resistance among the 95 spinach accessions (Table 1). The Verticillium disease showed a skewed distribution toward the susceptibility (Fig 1), indicting there were more susceptible spinach genotypes.

Genetic diversity and population structure

The population structure of the 95 spinach accessions was initially inferred using STRUCTURE 2.3.4 (Pritchard et al. 2000) and the peak of delta K was observed at K=2, indicating that the presence of the two main populations (clusters, Q1 and Q2) in the 95 accessions (Fig 2A and 2B). The classification of accessions into populations based on the model-based structure from STRUCTURE 2.3.4 was shown in Fig 2B and Table 1. In total 87 accessions (91.6%) were assigned to one of the two populations (Q1 and Q2). Population 1 and 2 (Q1 and Q2) consisted of 52 (54.7%) and 35 (36.8%) accessions, respectively. The remaining 8 accessions (8.4%) were categorized as having admixed ancestry between Q1 and Q2 called Q1Q2 (Table 1).

The genetic diversity among spinach accessions or cultivars was also assessed using the Maximum Likelihood (ML) method by MEGA 6 (Tamura et al. 2013). Several phylogenetic trees were drawn based on interpretation of results. We defined Q1 and Q2 as the two main clusters and used the same colors as the population structure Q1 (red) and Q2 (green) from the STRUCTURE 2.3.4 (Fig 2B) to draw the subtrees of the phylogenetic tree (Fig 2C) with Q1 (red and round shape), Q2 (green and square shape), and the admixture Q1Q2 (black empty square). Four phylogenetic trees were included: (1) Fig 2C, without taxon names in order to compare it to the structure populations and view them easily and clearly; (2) Supplementary Fig S1-1, the circle phylogenetic tree; and (3) Supplementary Figs S1-2 and S1-3: traditional rectangular phylogenetic trees, which are same as Fig S1-1 but different format. The phylogenetic trees from MEGA 6 (Fig 2C and Supplementary Fig S1-1, S1-2 and S1-3), were consistent with the structure populations (Q1-Q2) from STRUCTURE 2.3.4 (Fig 2A and 2B), indicating that there were two well-differentiated genetic populations and admixtures in the spinach panel.

Association analysis

In this research, we used three software programs and five models to do association analysis of Verticillium wilt resistance in order to identify the right SNP markers because different software programs and different models usually give different results. Three steps were used to identify SNP markers strongly associated with Verticillium wilt resistance. In the first step, we used an LOD value (LOD =~ (-LOG(P), where P is the P value estimated from TASSEL or GAPIT) having equaled or greater than 2.5 as the value from the all five models to screen SNP markers. If the LOD was 2.5 or

Table 1. Spinach accession, origin, Veriticillium wilt (VW) disease score, and cluster assigned in this study.

Accession	Original	VW scale	Cluster	Accession	Original	VW scale	Cluster
PI207518	Afghanistan	2.3	Q1	PI174389	Turkey	1.8	Q2
PI179588	Belgium	0.7	Q1	PI175923	Turkey	2	Q1
PI179589a	Belgium	2.6	Q2	PI175924	Turkey	3	Q1
PI179589b	Belgium	2.6	Q2	PI175925	Turkey	2.4	Q1
PI179590	Belgium	2	Q1	PI175927	Turkey	2.7	Q2
PI179591	Belgium	2	Q1	PI175929	Turkey	1.8	Q2
PI179592	Belgium	3	Q2	PI175931	Turkey	1.8	Q1
PI179593a	Belgium	2	Q1	PI176769	Turkey	2.3	Q1
PI179593b	Belgium	2	Q1	PI176772	Turkey	2	Q2
PI179594	Belgium	2	Q1	PI176774	Turkey	0.3	Q1
PI179597	Belgium	2	Q1	PI176775	Turkey	3	Q1
PI192945	China	3	Q1	PI176776	Turkey	2.5	Q1Q2
PI166366	India	3	Q1	PI176777	Turkey	2.2	Q1Q2
PI174960	India	1	Q1	PI176778	Turkey	2	Q2
PI175311	India	2.5	Q2	PI176779	Turkey	2	Q2
PI175312	India	3	Q1	PI177081	Turkey	2.3	Q1
PI175313	India	2.3	Q2	PI177557	Turkey	2	Q1
PI163310	Pakistan	3	Q1	PI177558	Turkey	2.3	Q1
PI184137	Serbia	3	Q1Q2	PI179041	Turkey	2.5	Q1Q2
PI206007	Sweden	2	Q2	PI179043	Turkey	3	Q1
PI179507	Syria	2	Q2	PI179044	Turkey	2.6	Q1
PI181809	Syria	2.5	Q2	PI204632	Turkey	1.6	Q1
PI181923	Syria	1.5	Q1	PI204732	Turkey	1	Q2
PI169673	Turkey	1.1	Q1	PI204733	Turkey	1.3	Q2
PI169675	Turkey	1	Q1	PI204734	Turkey	1.7	Q2
PI169676	Turkey	2	Q1	PI204735	Turkey	1	Q1
PI169677	Turkey	3	Q2	PI205231	Turkey	3	Q1
PI169679	Turkey	2.5	Q2	PI205232	Turkey	2.5	Q2
PI169682	Turkey	2	Q1	PI205234	Turkey	1.5	Q2
PI169683	Turkey	3	Q2	PI205235	Turkey	2.5	Q1
PI169684	Turkey	2.7	Q1	PI206473	Turkey	2.5	Q1
PI169690	Turkey	2	Q1	PI206474	Turkey	2.5	Q2
PI171858	Turkey	2.3	Q1Q2	PI206475	Turkey	2.4	Q2
PI171859	Turkey	1.2	Q1	PI206753	Turkey	2.5	Q1
PI171860	Turkey	2.2	Q2	NSL184380	US.California	2.5	Q1
PI171861	Turkey	1.2	Q2	NSL6084	US.California	2.5	Q1
PI171862	Turkey	1.5	Q1	NSL6086	US.California	2	Q1Q2
PI171863	Turkey	2	Q1	NSL81328	US.Maryland	0.5	Q1
PI171864	Turkey	2.2	Q2	PI169686	US.Maryland	1.5	Q2
PI173123	Turkey	2.5	Q1	NSL26513	US.Michigan	2.5	Q1
PI173127	Turkey	2.8	Q1	NSL40592	US.Michigan	1.5	Q1Q2
PI173128	Turkey	2.5	Q2	NSL6089	US.Missouri	1.7	Q2
PI173129	Turkey	2	Q2	NSL6095	US.Missouri	1.9	Q1
PI173130	Turkey	1.5	Q1Q2	NSL6082	US.NewYork	2.3	Q2
PI173131	Turkey	3	Q2	NSL6087	US.NewYork	1.7	Q2
PI174383	Turkey	2	Q1	NSL92513	US.Oregon	1	Q1
PI174387	Turkey	2	Q2	NSL6098	US.Virginia	3	Q1
PI174388	Turkey	2.4	Q1				

[a]VW scale = Verticillium wilt disease rating (0 – 4 scales) with LSM (least squared mean) calculated by ANOVA using JMP Genomics. [b]Cluster: Q1, Q2, and Q1Q2 are structured populations based on model-based populations in the verticillium wilt resistant association panel postulated by program STRUCTURE.

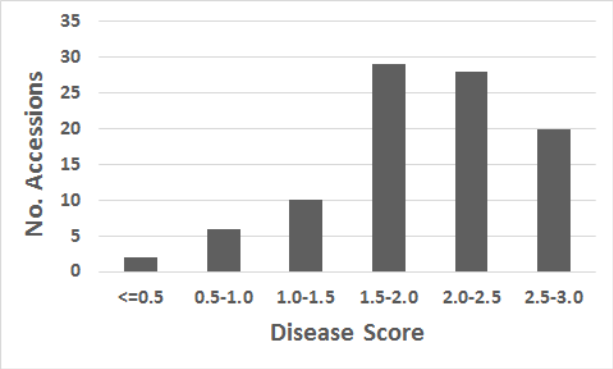

Fig 1. The distribution of Verticillium wilt disease rating in 95 spinach accessions.

above in one of the five models, the SNP was selected as the candidate marker. There were a total of 28 SNPs selected based on this threshold and their LOD values, P values, R-square percentages were listed in Supplementary Table S1. There were 15 SNPs from SMR analysis by QGene; 20 SNPs from GLM and 12 SNPs from MLM by TASSEL; and 12 SNPs from cMLM and 15 SNPs from EcMLM by GAPIT, which met the value with LOD 2.5 or higher (Table S1), indicating that the different software and different models gave us different results.

In the second step, we supposed that if it gave significant association in different models from different software, the SNP marker should be a good one. In this step, we used a LOD value having equal or greater than 2.0 across five models as the value to further identify the SNP markers associated with the Verticillium wilt resistance. There were 13 SNPs with LOD values equaled or greater than 2.0 across five models (Table 2). Among the 13 SNPs, seven SNPs, (AYZV02052660_2183, AYZV02064249_10266, AYZV02-112284_14543, AYZV02123399_146, AYZV02145765_ 3277, AYZV02199578_156, and AYZV02278250_41) had the LOD value equaled or greater than 2.5 across five models, indicating that they may be strongly associated with Verticillium wilt resistance.

In the last step, we selected the SNP markers based on the allele frequency for each SNP. For each SNP among the association panel – there are 95 spinach genotypes in this study, we selected the SNP having both allelic homogeneity A and B with less heterogeneity H and less missing data. Among the 13 SNPs selected from the second step (Table 2), eight SNPs, AYZV02052660_2183, AYZV02064249 _10266, AYZV02145765_3277, AYZV02193759_955, AYZV02194583_23976, AYZV02199578_156, AYZV02278250_41, and AYZV02297667_1112 had no spinach genotype in one of two allelic homogeneities either A or B. Therefore, we think that the eight SNPs are not good SNP markers. The five SNPs left, AYZV02052595_108, AYZV02112284_14543, AYZV02123399_146, AYZV0216-4612_331, and AYZV02170942_274 had both allelic homogeneities A and B with less heterogeneity H and they were identified to be the SNP markers associated with Verticillium wilt resistance from this study (Table 2).

Among the five SNP markers, AYZV02112284_14543 had LOD values equal or greater than 2.5 across five models with high R-squared values from 9.3 to 18.2%, suggesting that it was strongly associated with Verticillium wilt resistance. Other four SNPs also had high R-squared values from 9.3 to 17.2%, explaining large variations of the disease resistance (Table 2), suggesting the five SNP markers to be associated with the Verticillium wilt resistance.

Discussion

Two pathogenic races of *V. dahliae* have been identified in lettuce and tomato (Baergen et al., 1993; Hayes et al., 2011; Vallad et al., 2006), and both of these races having different isolates can be recovered from infested spinach (*Spinacia oleracea* L.) seeds (Mou et al., 2015; Short et al., 2014). A race-specific PCR assay identified 96% of isolates as race 2 from among the 340 *V. dahliae* isolates recovered from spinach seeds produced in Chile, Denmark, the Netherlands, and the United States (Short et al., 2014). With the deployment of race 1 resistance in lettuce cultivars (Hayes et al., 2011), it is expected that the proportion of race 2 pathogen will increase in California soils under the genetic selection pressure. For this reason, identification and development of spinach germplasm with resistance to race 2 isolates is needed. Therefore, we used a race 2 isolate (So 923) of *V. dahliae* from spinach to screen the spinach germplasm in this study.

From this research, significant genetic variance of Verticillium wilt disease resistance was observed among the 95 spinach accessions with a wide range of disease severities observed from 0.0 to 3.0 out of the 0-4 rating scale, which was similar to the results reported by Villarroel-Zeballos et al. (2012). There were more susceptible spinach accessions in the tested spinach panel. Villarroel-Zeballos et al. (2012) also reported that there were more susceptible accessions in the120 spinach accessions and 10 commercial hybrids tested. Although there were many susceptible genotypes, some resistant accessions were observed in both studies. In this research, eight spinach genotypes, NSL 81328, NSL 92513, PI 169675, PI 174960, PI 176774, PI 179588, PI 204732, and PI 204735 showed the highest Verticillium wilt resistance with disease severity ratings of 1.0 or less (Table 1), indicating that the eight spinach genotypes could be used as parents in spinach breeding programs to improve Verticillium wilt resistance. Villarroel-Zeballos et al. (2012) also observed that five spinach accessions, PI 175931, Ames 26243, PI 163309, PI 261789, and PI 494751 showed consistently low disease severity ratings in three experiments with and without inoculum of *V. dahliae* pathogen. Although 120 spinach accessions and 10 commercial hybrids were used in the study by Villarroel-Zeballos et al. (2012), only the most resistant and susceptible spinach accessions and commercial spinach hybrids were reported. Of these spinach genotypes, only one accession, PI175931 was included in this research and it showed intermediate resistance with 1.8 disease rating of resistance to the a race 2 isolate. PI175931 was ranked number 23 out of 95 tested spinach accessions in this study, but it was ranked number 30 with disease severity 21.9%, number 1 with disease severity 5.7%, and number 6 with disease severity 8.3% in trial 1 (2006), trial 2 (2007) and trial 3 (2008) out of 130 spinach accessions in the study by Villarroel-Zeballos et al. (2012). Both studies showed that the PI175931 had partial resistance to *V. dahliae*, however, the difference may be caused by the different isolates used and experimental conditions. Beiquan et al. (2015) screened 268 USDA spinach germplasm accessions plus 9 commercial cultivars for Verticillium wilt resistance with two isolates So 923 and So 925 of *V. dahliae* race 2 and one isolate So 302 of race 1. From 18 accessions plus two cultivars reported, there were a wide range of disease incidence and severity and different responses were observed to different races or isolates, i.e. some spinach accessions showed resistance to race 1 but were susceptible to race 2, or conversely, even to different isolates, So 923 and So 925 of the same race 2.

Besides the eight accessions with disease severity ratings of 1.0 or less, other four accessions also showed mediate resistance with disease severity 1.3 or less (Supplementary Table S1). The genetic relationships among the 12 accessions were also analyzed using 2878 SNPs by MEGA 6 and a phylogenetic tree was drawn (Fig 3), in which each spinach genotype was drawn using a accession_ID combining the accession number, the original country (region), structured populations (Q1 and Q2, or mixture Q1Q2), and the scale of Verticillium wilt disease severity in order to view the information easily from the tree. From the phylogenetic tree, three Q2 accessions merged together, further validated the population structured. The two NSL accessions, NSL81328 and NSL92513, originally from US Maryland and Oregon merged together, indicating that they had similar genetic

Table 2. Thirteen SNP markers associated with Verticillium wilt resistance identified from five models using QGene, Tassel and GAPIT in 95 spinach accessions.

SNP name[a]	LOD (-Log(P)) value					R-square value (%)				
	SMR[b]	GLM[b]	MLM[b]	cMLM[b]	EcMLM[b]	SMR	GLM	MLM	cMLM	EcMLM
AYZV02052595_108	3.1	2.5	2.3	2.1	2.6	13.8	10.5	11.6	13.4	14.6
AYZV02112284_14543	4.1	2.9	2.5	2.8	3.1	18.2	9.7	9.3	17.0	17.0
AYZV02123399_146	3.4	3.1	2.8	2.7	3.1	15	13.4	14.0	16.4	17.2
AYZV02164612_331	3.4	2.1	2.0	2.1	2.4	15.1	9.3	9.5	13.4	13.6
AYZV02170942_274	2.7	2.5	2.3	2.2	2.1	12.5	10.2	10.9	13.8	12.2
AYZV02052660_2183	2.9	2.8	2.6	2.6	2.6	13.3	9.6	10.1	16.1	14.5
AYZV02064249_10266	3.1	2.8	2.6	2.9	3.1	13.8	9.4	9.6	17.3	17.0
AYZV02145765_3277	3.5	3.2	3.0	2.7	3.0	15.7	11.1	11.4	16.6	16.7
AYZV02193759_955	2.1	2.2	2.2	2.8	2.5	9.9	6.7	7.6	16.7	14.3
AYZV02194583_23976	2.4	3.0	2.9	2.2	2.8	11	10.1	10.9	14.1	15.8
AYZV02199578_156	3.3	3.1	2.8	2.9	3.0	14.8	10.2	10.2	17.2	16.7
AYZV02278250_41	2.5	3.0	2.9	2.6	2.8	11.4	10.3	11.7	16.1	15.8
AYZV02297667_1112	2.4	2.7	2.4	2.5	2.9	11.1	8.7	9.0	15.3	16.0
Spinach genome Spinach-1.0.3 information				Viroflay-1.0.1		SNP allele frequency (%)[c]				MAF (%)[d]
SNP name	SNP Type	Contig at AYZV02 project	SNP Position	Contig at AYZV01 project	SNP Position	A	B	H	Missing	
AYZV02052595_108	T/C	AYZV02052595	108	AYZV01040503	108	92.6	3.2	1.1	3.2	3.8
AYZV02112284_14543	C/T	AYZV02112284	14543	AYZV01084682	14543	85.3	10.5	0.0	4.2	11.0
AYZV02123399_146	G/A	AYZV02123399	146	AYZV01092619	146	2.1	85.3	8.4	4.2	6.6
AYZV02164612_331	A/T	AYZV02164612	331	AYZV01121369	331	77.9	12.6	5.3	4.2	15.9
AYZV02170942_274	G/T	AYZV02170942	274	AYZV01125597	274	94.7	2.1	3.2	0.0	3.7
AYZV02052660_2183	A/G	AYZV02052660	2183	AYZV01040555	2183	72.6	0.0	21.1	6.3	11.2
AYZV02064249_10266	T/G	AYZV02064249	10266	AYZV01049272	10266	0.0	89.5	7.4	3.2	3.8
AYZV02145765_3277	T/G	AYZV02145765	3277	AYZV01108434	3277	0.0	78.9	15.8	5.3	8.3
AYZV02193759_955	T/A	AYZV02193759	955	AYZV01141173	955	0.0	84.2	15.8	0.0	7.9
AYZV02194583_23976	G/A	AYZV02194583	23976	AYZV01141714	23976	0.0	90.5	6.3	3.2	3.3
AYZV02199578_156	G/A	AYZV02199578	156	AYZV01145092	156	0.0	90.5	8.4	1.1	4.3
AYZV02278250_41	C/A	AYZV02278250	41	AYZV01203883	41	0.0	85.3	8.4	6.3	4.5
AYZV02297667_1112	G/A	AYZV02297667	1112	AYZV01222189	1112	0.0	63.2	33.7	3.2	17.4

[a]SNP name is defined as the contig name plus the SNP position on the contig. [b]SMR = single marker regression using the QGene 4.3.10 (Joehanes & Nelson 2008); GLM = regression linear model; and MLM = mixed linear model using TASSEL 5 (Bradbury et al. 2007; http://www.maizegenetics.net/tassel); and cMLM = compressed mixed linear model and EcMLM = enriched compressed mixed linear model methods using GAPIT (Li et al. 2014; Lipka et al. 2012; Zhange et al. 2010). [c]SNP allele frequency (%) represents each SNP with its homogenous allele A and B, heterogenous allele H, and its missing percentage among the 95 spinach samples. [d]MAF = minor allele frequency for each SNP.

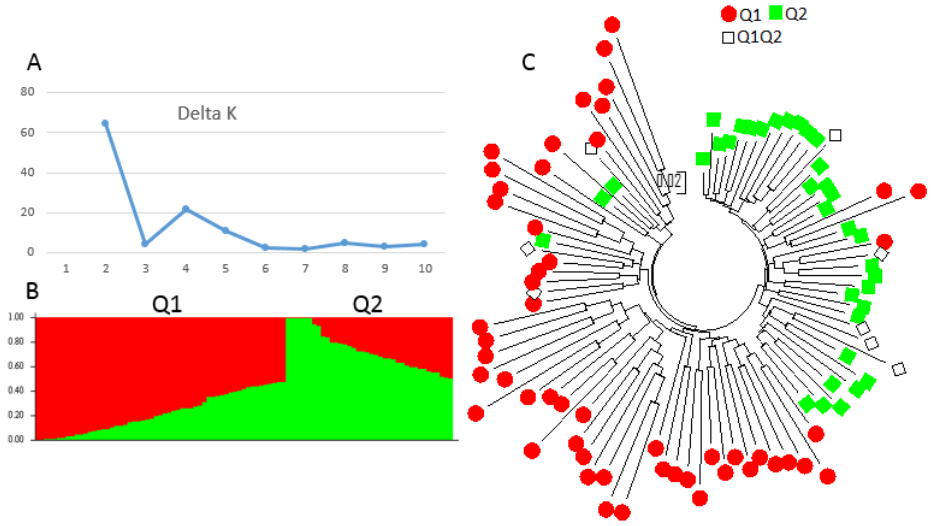

Fig 2. Model-based populations in the Verticillium wilt resistant association panel (A) Delta K values for different numbers of populations assumed (K) in the STRUCTURE analysis (B) Classification of 95 spinach accessions into two populations Q1 (red) and Q2 (green) using STRUCTURE 2.3.4. The distribution of the accessions to different populations is indicated by the color code (Q1: red and round shape, Q2: green and square shape, Q1Q2: black empty square; (C) Maximum Likelihood (ML) tree of the 95 accessions drawn by MEGA 6. The color codes for each population and admixture are consistent in the figure B and C.

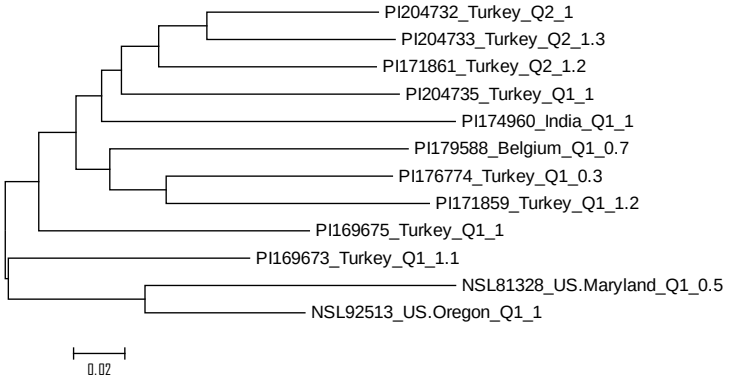

Fig 3. The phylogenetic tree among 12 Verticillium wilt resistant spinach accessions using 2878 SNPs by MEGA 6, in which each spinach genotype was drawn using a accession_ID combining the accession number, the original country (region), structured populations (Q1 and Q2, or mixture Q1Q2), and the scale of Verticillium wilt disease severity in order to view the information easily from the tree.

backgrounds caused by geographic effect. Besides the two Turkey accessions, PI69673 and PI69675, that were separate from others, the eight accessions left merged together to form a group, indicating that they had closer genetic distances and suggesting similarity of genetic backgrounds among them. In order to increase genetic diversity and variance in breeding programs, the spinach germplasm with different backgrounds will be selected as parents. The phylogenetic analysis will provide breeders with useful information on how to use these Verticillium wilt resistant germplasm. So far, although the genetics of Verticillium wilt resistance in spinach is still not characterized in terms of whether the resistance is controlled by major genes or minor genes through genetic study, the nature of quantitative resistance of *V. dahlia* was postulated (Villarroel-Zeballos et al., 2012). Villarroel-Zeballos et al. (2012) concluded that it was unlikely that qualitative or major resistance to *V. dahliae* was present. Because no single SNP showed very high LOD value, we also propose that the Verticillium wilt resistance is a quantitative trait controlled by several minor genes. Whether there are major resistance genes in spinach for Verticillium wilt resistance needs further study.

In this study, with three software programs and five models to conduct association analyses of Verticillium wilt resistance, we observed a lot of SNPs that showed different results in different models from different software programs. We use three steps to identify SNP markers to be strongly associated with Verticillium wilt resistance. In the first step, we used an LOD value equal or greater than 2.5 in one of the five models, 28 SNPs were selected for further consideration. In the second step, we used a LOD value equal or greater than 2.0 across five models to select SNP markers and supposed that if it gave significant association in different models from different softwares, the SNP marker should be a good one, and 13 SNPs were selected for further identification. In the last step, we selected SNP markers based on SNP allelic homogeneity rate and finally five SNP markers were recognized as the SNP markers identified from this study. We know this is not a perfect approach to identify SNP markers, but it seems a reasonable one to select SNP markers in this study.

The SNP type was [C/T], [A/T], [G/T], [G/A], and [T/C] for AYZV02112284_14543, AYZV02164612_331, AYZV0217-0942_274, AYZV02123399_146, and AYZV02052595_108, respectively (Table 2). The beneficial allele, i.e. the allele with low Verticillium wilt disease severity in each SNP, was T, T, T, A, and C, respectively for the five SNP markers, and a T-test showed that there was significant difference between two alleles in each SNP with P value less than 0.001, suggesting these markers are effective to be used to select resistant genotypes. It is interesting that the PI176774 accession with the 0.3 lowest severity rating had five beneficial alleles among the five SNP markers. After all, the five SNP markers identified from this study had LOD values equal or greater than 2.0 across five models and also had high R-squared values from 9.3 to 18.2%. It is expected that these SNP markers provide a tool utilized in spinach molecular breeding efforts to select Verticillium wilt resistant plants in a breeding program.

Materials and Methods

Plant materials

A total of 95 accessions of spinach (*Spinacia oleracea*) USDA-GRIN germplasm originally collected from 10 countries were used for association analysis of Verticillium wilt resistance in spinach (Table 1). All seeds were kindly provided by the North Central Regional Plant Introduction Station, USDA-ARS, Iowa State University, Ames, Iowa, USA.

Disease evaluation

The experiment was conducted under greenhouse conditions at Salinas, California, US, described by Mou et al. (2015). The experimental design was a randomized complete block (RCBD) with three replications. In each replication, eight seeds of each accession were planted in Sunshine Plug 5 Growing Mix (Sun Gro Horticulture, Agawam, MA) in plastic transplanting trays (128 cells, 3 x 3 x 5 cm in length x width x height) in a greenhouse in winter to control day length. The three replications were inoculated with a Race 2 isolate So 923 of *V. dahlia* from spinach (Atallah et al., 2011). Seedlings were inoculated at 3, 4, and 5 weeks after sowing by saturating the soil in each plug tray well with a 3-ml suspension containing 2 x 10^6 conidia/ml in sterile, distilled water. Seedlings were incubated for another week after the last then transplanted into 0.5-liter (16 oz) foam-insulated cups filled with a mixture of pasteurized sand: Sunshine Plug 5 Growing Mix (3:1, vol/vol). One week after transplanting, day length was extended to 19 hr/day by supplemental lighting to promote bolting, as symptoms of Verticillium wilt on spinach developed after the bolting stage (Mou et al., 2015). The *V. dahliae* inoculation and disease evaluation described by Mou et al. (2015) were used. Beginning three weeks after the last inoculation, severity of symptoms were rated weekly using a scale of 0 to 4: 0 = no symptoms, 1 = lower leaves with patches of yellow areas or wilting, 2 = middle leaves with patches of yellow areas or wilting, 3 = upper leaves with patches of yellow areas or wilting, and 4 = all leaves died. After final rating, roots were cleaned of sand and cut longitudinally to evaluate the disease as the % brown discoloration of vascular tissue in the roots, crown, and lower stem, characteristic of Verticillium wilt. The growth period (from planting to death of all leaves) of the inoculated plants was compared with the uninoculated control. To confirm the presence of the pathogen, *V. dahliae* was re-isolated from diseased tissue. Roots, crown, and lower stems were placed on NP-10 medium after surface sterilization (1% bleach solution for 1 min) and examined microscopically for development of conidiophores and/or microsclerotia of *V. dahliae*. To examine the seed transmission of the pathogen, mature seeds from each plant were harvested separately and assayed for *V. dahliae* by plating 20 seeds on NP-10 medium (Kabir et al., 2004; Short et al., 2014). The seeds were observed under a microscope for microsclerotia and/or conidiophores and conidial characteristics of *V. dahliae*. Phenotypic data of the final Verticillium wilt disease rating were analyzed using Microsoft (MS) Excel 2013 for the average, range, standard deviation (Stdev), and Stdev error. The distributions of Verticillium wilt was also drawn using MS Excel. The Verticillium wilt scores were analyzed with ANOVA (Analysis of Variance) using JMP Genomics 7 software (SAS Institute, Cary, NC, USA) with spinach genotypes as fixed model. The least squared mean (LSM) of each spinach genotype from JMP was used as the phenotypic data in the association mapping of Verticillium wilt resistance.

DNA extraction, GBS and SNP discovery

Genomic DNA was extracted from dried leaves of spinach plants using the CTAB (hexadecyltrimethyl ammonium bromide) method (Kisha et al., 1997). Before DNA extraction, the fresh leaves of each sample were freeze dried at -20C for 48 hours in a lab lyophilizer. DNA library was prepared using the restriction enzyme *Ape*KI following the GBS protocol described by Elshire et al. (2011). The 90 bp double-end sequencing was performed on each spinach genotype using GBS protocol by an Illumina HiSeq 2000 at BGI Genomics Research Institute (BGI) in Hong Kong. GBS data assembly, mapping and SNP discovery were done using SOAP family software (http://soap.genomics.org.cn/) by the bioinformatics team in BGI. The GBS data provided by BGI averaged 3.26 M with 90 bp short-read, and 283.74 Mbp nucleotides for spinach samples. The short reads of the GBS data were aligned to spinach genome reference AYZV01 and AYZV02 (http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV01 and http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV02) using SOAPaligner/soap2 (http://soap.genomics.org.cn/) and SOAPsnp v 1.05 was used for SNP calling by the bioinformatics team at BGI (Li, 2011; Li et al., 2009). Approximately a half million SNPs were discovered from the GBS data among the 95 spinach germplasm accessions and the original SNP data were also provided by BGI. The spinach accessions and SNPs were filtered before conducting genetic diversity and association analyses. The SNP data were filtered by minor allele frequency (MAF) > 2%, missing data < 15%, and heterozygous genotype < 35%, where MAF refers to the frequency at which the least common allele occurs in a given population. In human International HapMap Project, SNPs with 5% or greater MAF were targeted by the HapMap project (The International HapMap Consortium, 2005). After filtering, 2,878 SNPs among 95 spinach accessions were used for genetic diversity and association analysis in this study. Although the 2,878 SNP data across 95 spinach genotypes were not enclosed, the 28 SNP data which were identified to be candidate markers for Verticillium wilt resistance were listed in supplementary Table S2.

Genetic diversity and population structure analysis

A model-based clustering method in STRUCTURE 2.3.4 program (Pritchard et al. 2000) was used to infer population

structure. In order to identify the number of populations (K) capturing the major structure in the data, the burn-in period was set at 50,000 with the Markov Chain Monte Carlo iterations and the run length was set at 10,000 in an admixture model and correlated allele frequencies that were independent for each run (Lv et al., 2012). Ten runs were performed for each simulated value of K, ranging from 1 to 10. The delta K was calculated using the formula described by Evanno et al. (2005). The optimal K was determined using Structure Harvester (Earl and vonHoldt 2012; http://taylor0.biology.ucla.edu/structureHarvester/). After the optimal K was determined, a Q-matrix was obtained and was used in Tassel 5 for association analysis of Verticillium wilt resistance. Each spinach accession was also signed to each cluster (Q) based on the probability of this germplasm in the clusters and the cut-off probability for assignment of an accession to a cluster was 0.50. Based on the optimal K, a Bar plot with 'Sort by Q' was obtained to show the visual of the population structure among the spinach association panel. Genetic diversity among spinach accessions or cultivars was also assessed and the phylogeny trees were drawn using MEGA 6 (Tamura et al., 2013) based on the Maximum Likelihood tree method with the following parameters. Test of Phylogeny: Bootstrap Method, No. of Bootstrap Replications: 500, Model/Method: General Time Reversible model, Rates among Sites: Gamma distributed with Invariant sites (G + I), Number of Discrete Gamma Categories: 5, Gaps/Missing Data Treatment: Use all sites, ML Heuristic Method: Subtree-Pruning-Regrafting-Extensive (SPR level 5), Initial Tree for ML: Make initial tree automatically (Neighbor Joining), and Branch Swap Filter: Moderate. During the drawing of the phylogeny trees, the population structure and the cluster information were imported to MEGA 6 for combined analysis of genetic diversity. For sub-tree of each Q (cluster), the shape of 'Node/Subtree Marker' and the 'Branch Line' was drawn with the same color as in the figure of the Bar plot of the population clusters from the STRUCTURE analysis.

Association analysis

Association analysis was conducted with the regression linear model (RLM) and the mixed linear model (MLM) methods as described in TASSEL 5 (Bradbury et al. 2007; http://www.maizegenetics.net/tassel) and the analysis was also performed with compressed mixed linear model (Zhang et al., 2010) and enriched compressed mixed linear model (EcMLM) (Li, et al., 2014) implemented in the GAPIT R package (Lipka et al., 2012). The QGene 4.3.10 was also used to conduct SMR for all SNPs (Joehanes & Nelson, 2008). Although QGene was developed for QTL mapping, it can also be used in association analysis through SMR.

Conclusions

A significant genetic variation of Verticillium wilt disease resistance was observed among the 95 spinach accessions tested in this study. Eight spinach genotypes, NSL 81328, NSL 92513, PI 169675, PI 174960, PI 176774, PI 179588, PI 204732, and PI 204735 were identified with highest Verticillium wilt resistance and they could be used as parents in spinach breeding programs to improve Verticillium wilt resistance. Two well-differentiated genetic populations and admixtures among the 95 accessions were postulated and provided the genetic diversity information for breeders to select these germplasm in breeding. Five SNP markers were identified to be strongly associated with Verticillium wilt resistance and they may provide a tool utilized in spinach molecular breeding efforts to select Verticillium wilt resistant plants in a breeding program.

Acknowledgments

This work is supported, in part, by the USDA National Institute of Food and Agriculture Hatch project accession number 1002423.

Reference

Atallah ZK, Maruthachalam K, Vallad GE, Davis RM, Klosterman SJ, Subbarao KV (2011) Analysis of *Verticillium dahliae* suggests a lack of correlation between genotypic diversity and virulence phenotypes. Plant Dis. 95: 1224-1232.

Baergen K, Hewitt J, Clair DS (1993) Resistance of tomato genotypes to four isolates of *Verticillium dahliae* race 2. HortScience. 28: 833-836.

Bhat RG, Subbarao KV (1999) Host range specificity in *Verticillium dahliae*. Phytopathology. 89: 1218-1225.

Bryant D (2009) Verticillium wilt in lettuce and spinach. Western Farm Press. Available <http://westernfarmpress.com/vegetables/verticillium-wilt-lettuce-and-spinach>. Accessed in: May 2, 2016.

Collard BCY, Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil Trans R Soc B12. 363: 557-572.

Correll JC, Bluhm BH, Feng C, Lamour K, du Toit LJ, Koike ST (2011) Spinach: Better management of downy mildew and white rust through genomics. Eur J Plant Pathol. 129: 193-205.

Correll JC, Morelock TE, Black MC, Koike ST, Brandenberger LP, Dainello FJ (1994) Economically important diseases of spinach. Plant Dis. 78: 653-660.

Dicoteau DR (2000) Vegetable crops. New Jersey, Printice Hall.

Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutierrez S, Zakrzewski F, Tafer H, Rupp O, Sörensen TR, Stracke R, Reinhardt R, Goesmann A, Kraft T, Schulz B, Stadler PF, Schmidt T, Gabaldon T, Lehrach H, Weisshaar B, Himmelbauer H (2014) The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). Nature. 505: 546–549.

du Toit LJ, Derie ML, Hernandez-Perez P (2005). Verticillium wilt in spinach seed production. Plant Dis. 89: 4-11.

Duressa D, Rauscher G, Koike ST, Mou B., Hayes RJ, Maruthachalam K, Subbarao KV, Klosterman SJ (2012) A realtime PCR assay for detection and quantification of *Verticillium dahliae* in spinach seed. Phytopathology. 102: 443-451.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 4 (2): 359-361.

Elshire RJ, Glaubitz JC, Sun Q (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One. 6(5): e19379.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14: 2611–2620.

Hayes RJ, Maruthachalam K, Vallad GE, Klosterman SJ, Subbarao KV (2011) Selection for resistance to Verticillium wilt caused by race 2 isolates of *Verticillium*

dahliae in accessions of lettuce (*Lactuca sativa* L.). HortScience. 46: 201-206.

He J, Zhao X, Laroche A, Lu Z, Liu H, Li Z (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front Plant Sci. 5: 484. doi: 10.3389/fpls.2014.00484.

The International HapMap Consortium (2005) A haplotype map of the human genome. Nature. 437: 1299–1320.

Iglesias-Garcia AM, Villarroel-Zeballos MI, Feng C, du Toit LJ, Correll JC (2013) Pathogenicity, virulence, and vegetative compatibility grouping of Verticillium isolates from spinach seed. Plant Dis. 97: 1457-1469.

Joehanes R, Nelson JC (2008) QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics. 24: 2788-2789.

Kabir Z, Bhat RG, Subbarao KV (2004) Comparison of media for recovery of *Verticillium dahliae* from soil. Plant Dis. 88: 49-55.

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

Koike ST, Cahn M, Cantwell M, Fennimore S, Natwick E, Smith RF, Takele E (2011) Spinach production in California. Univ Calif Agric Nat Resour Publ. 7212.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 27: 2987-2993.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J (2009) The sequence alignment/ map format and SAMtools. Bioinformatics. 25: 2078-2079.

Li M, Li X, Bradbury P, Yu J, Zhang Y, Todhunter RJ, Buckler ES, Zhang Z (2014) Enrichment of statistical power for genome-wide association studies. BMC Biology. 12:73.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbur PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics. 28: 2397–2399.

Lv J, Qi J, Shi Q, Shen D, Zhang S, Zhang A, Shao G, Li H, Sun Z, Weng Y, Shang Y, Gu X, Li X, Zhu X, Zhang J, van Treuren R, van Dooijeweert W, Zhang Z, Huang S (2012) Genetic diversity and population STRUCTURE of cucumber (*Cucumis sativus* L). PLoS ONE 7(10): e46919

Maruthachalam K, Klosterman SJ, Anchieta AG, Mou B, Subbarao KV (2013) Colonization of spinach by *Verticillium dahliae* and effects of pathogen localization on the efficacy of seed treatments. Phytopathology. 103: 268-280.

Minoche AE et al (2015) Exploiting single-molecule transcript sequencing for eukaryotic gene prediction. Genome Biol. 16: 184.

Morelock TE, Correll JC (2008) Spinach. In: Prohens J and Nuez F, Eds., Handbook of Plant Breeding, Vegetables I, *Asteraceae*, *Brassicaceae*, *Chenopodicaceae*, and *Cucurbitaceae*. Springer, New York. pp. 189-218.

Mou B, Klosterman SJ, Anchieta AG, Wood EM, Subbarao K (2015) Characterization of spinach germplasm for resistance against two races of *Verticillium dahliae*. HortScience. 50: 1631-1635.

Poland JA, Rife TW (2012).Genotyping-by-sequencing for plant breeding and genetics. Plant Genome. 5: 92–102.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics. 155: 945-959.

Sackston WE, Sedun FS (1982) Verticillium wilt of spinach: A useful experimental system. Can J Plant Pathol. 4: 310.

Shi A, Buckley B, Mou B, Motes D, Morris JB, Ma J, Xiong H, Qin J, Yang W, Chitwood J, Weng Y, Lu W (2016) Association analysis of cowpea bacterial blight resistance in USDA cowpea germplasm. Euphytica. 208: 143–155.

Short DP, Gurung S, Koike ST, Klosterman SJ, Subbarao KV (2015) Frequency of Verticillium species in commercial spinach fields and transmission of *V. dahliae* from spinach to subsequent lettuce crops. Phytopathology. 105: 80-90.

Short DP, Gurung S, Maruthachalam K, Atallah ZK, Subbarao KV (2014) *Verticillium dahliae* race 2-specific PCR reveals a high frequency of race 2 strains in commercial spinach seed lots and delineates race structure. Phytopathology. 104: 779-785.

Sonah H, Bastien M, Iquira E, Tardivel A, Legare G (2013) An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PLoS ONE. 8(1): e54603.

Tamura K, Stecher G, Peterson D, Ailipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30: 2725-2729.

Vallad GE, Qin QM, Grube RC, Hayes RJ, Subbarao KV (2006) Characterization of race-specific interactions among isolates of *Verticillium dahliae* pathogenic on lettuce. Phytopathology. 96: 1380-1387.

van Deynze A (2015) A De Novo Draft Assembly of Spinach Using Pacific Biosciences Technology. Plant & Animal Genomics XXII conference, January 10-15, 2014, San Diego, CA. Available in: http://aa314.gondor.co/webinar/a-de-novo-draft-assembly-of-spinach-using-pacific-biosciences-technology/. Accessed in: May 2, 2016.

van Deynze A, Ashrafi H, Hickey L, Peluso P, Rank D, Chin J, Rapicavoli N, Drake J, Garvin T, Schatz M (2015) Using spinach to compare technologies for whole genome assemblies. Plant & Animal Genomics XXIII conference, January 10-14, 2015, San Diego, CA.

van der Spek J (1972) Internal carriage of *Verticillium dahliae* by seeds and its consequence. Meded Fac Landbouwwet Rijksuniv. Genetics. 37: 567-573.

Villarroel-Zeballos MI, Feng C, Iglesias A, du Toit LJ, Correll JC (2012) Screening for resistance to *Verticillium dahliae* in spinach and isolation of *Verticillium dahliae* from seed of spinach accessions. HortScience. 47(9): 1297-1303.

Xu Y, Crouch JH (2008) Marker-assisted selection in plant breeding: from publications to practice. Crop Sci. 48: 391–407.

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 42: 355-360.