HortScience 51(5):481–486. 2016.

Population Structure and Association Analysis of Bolting, Plant Height, and Leaf Erectness in Spinach

Jessica Chitwood and Ainong Shi[1]
Department of Horticulture, University of Arkansas, 316 PTSC, Fayetteville, AR 72701

Beiquan Mou[1]
U.S. Department of Agriculture, Agricultural Research Service, 1636 East Alisal Street, Salinas, CA 93905

Michael Evans and John Clark
Department of Horticulture, University of Arkansas, 316 PTSC, Fayetteville, AR 72701

Dennis Motes
Vegetable Research Center, University of Arkansas, Alma, AR 72921

Pengyin Chen
Department of Crop, Soil, and Environmental Sciences, 215 PTSC, University of Arkansas, Fayetteville, AR 72701

David Hensley
Department of Horticulture, University of Arkansas, 316 PTSC, Fayetteville, AR 72701

Additional index words. Spinacia oleracea, association mapping, single nucleotide polymorphism (SNP), genotyping by sequencing (GBS)

Abstract. **Spinach (*Spinacia oleracea* L.) is an important vegetable worldwide with high nutritional and health-promoting compounds. Bolting is an important trait to consider to grow spinach in different seasons and regions. Plant height and leaf erectness are important traits for machine harvesting. Breeding slow bolting, taller, and more erect spinach cultivars is needed for improved spinach production. A total of 288 United States Department of Agriculture (USDA) spinach accessions were used as the association panel in this research. Single-nucleotide polymorphisms (SNPs) discovered through genotyping by sequencing (GBS) were used for genotyping. Two structured populations and the admixtures were inferred for the 288 spinach accession panel using STRUCTURE and MEGA. Association mapping was conducted using single-marker regression (SMR) in QGene, and general linear model (GLM) and mixed linear model (MLM) built in TASSEL. Three SNP markers, AYZV02001321_398, AYZV02041012_1060, and AYZV02118171_95 were identified to be associated with bolting. Eight SNP markers, AYZV02014270_540, AYZV02250508_2162, AYZV02091523_19842, AYZV02141794_376, AYZV02077023_64, AYZV02210662_2532, AYZV02153224_2197, and AYZV02003975_248 were found to be associated with plant height. Four SNP markers, AYZV02188832_229, AYZV02219088_79, AYZV02030116_256, and AYZV02129827_197 were associated with erectness. These SNP markers may provide breeders with a tool in spinach molecular breeding to select spinach bolting, plant height, and erectness through marker-assisted selection (MAS).**

Molecular markers have become of increasing importance in plant breeding. For many major crop species, potential genetic variation for important agronomic traits already exists with varying degrees of accessibility (Thomson et al., 2010). DNA markers for genes of interest allow breeders to make selections when otherwise the gene for the trait may have been masked by heterozygosity. Association mapping is a relatively recent technology development, which identifies quantitative trait loci (QTLs) associated with phenotypic characteristics (Zhu et al., 2008) and provides the link for breeders to make selections based on genetic information.

Molecular markers and MAS have been successfully used to select specific genes/alleles in plant breeding, and as cost decreases along with rapid improvement of the technology, these methods are becoming more widely used (Kumar et al., 2012; Morelock and Correll, 2008; Thomson et al., 2010). Genetic research across many disciplines, from human genomic studies to marker-assisted breeding of livestock and plants, uses SNPs as the marker of choice for various reasons, but especially their abundance within any genome and cost efficiency (Zhu et al., 2008). The use of SNPs has become a powerful tool for gaining a better understanding of plant genomics by mapping chromosomes via association mapping and tagging important genes, as well as diversity analysis and other studies (Kumar et al., 2012). Association mapping has been used to successfully identify markers and loci associated with major agronomic traits (Lakew et al., 2013) such as anthracnose resistance in sorghum (*Sorghum bicolor* L.) (Upadhyaya et al., 2013), growth habit and days to flowering in common bean (*Phaseolus vulgaris* L.) (Nemli et al., 2014), and heat tolerance in cowpea (Lucas et al., 2013).

Some of the major agronomic traits of interest in spinach are bolting, plant height, and leaf erectness. Bolting is an important trait to consider in relation to developing spinach cultivars for year-round production because of its sensitivity to photoperiod (Chun et al., 2000). Long-day exposure induces bolting in spinach, rendering the plant unmarketable (Goreta and Leskovar, 2006). Because some commercially grown spinach is cut multiple times (Morelock and Correll, 2008), overwintered spinach that is susceptible to bolting in the spring reduces the number of harvests that may be taken and therefore reduces overall yield. Genetic variation among spinach for bolting has been documented for many years, and therefore, late-bolting cultivars can be developed through breeding efforts (Goreta and Leskovar, 2006).

Commercial spinach cultivation is highly mechanized (Koike et al., 2011; Morelock and Correll, 2008), and traits such as plant height and erectness affect the ability to harvest the plants. Plant height in spinach is a complex trait and a range of phenotypic values often occur. Spinach erectness refers to how close to or far away from the ground the spinach leaves lie on a mature plant. In the United States, erect leaves are generally preferred to accommodate high-density spinach production and mechanical harvesting.

To date, knowledge of the spinach genome is limited and few reports have been published on the use of molecular markers in spinach. Khattak et al. (2006) published a genetic linkage map with six linkage groups, constructing the map with 101 amplified fragment length polymorphisms (AFLPs) and nine simple sequence repeat (SSRs). This genetic map has a total length of 585 cM, and with an average distance of 5.18 cM between markers (Khattak et al., 2006), but does not offer a great amount of detail about the linkage groups. AFLPs and SSRs, while useful, are less specific than SNP markers. Recently, Chan-Navarrete et al. (2016) first reported an SNP

Received for publication 7 Jan. 2016. Accepted for publication 28 Mar. 2016.
This work is supported, at least in part, by the USDA National Institute of Food and Agriculture Hatch project accession number 1002423.
[1]Corresponding author. E-mail: ashi@uark.edu or Beiquan.Mou@ars.usda.gov.

genetic maps of six linkage groups (P01-P06) consisted of 283 SNP markers, ranging in size from 46 to 116 cm and identified 39 QTLs related to nitrogen use efficiency in spinach. The identification of SNP markers for spinach traits of interest, including bolting, plant height, and erectness, will provide breeders with powerful tools to develop improved spinach cultivars more efficiently. Therefore, the development of robust SNP markers and SNP genetic maps would be a valuable resource for spinach breeding efforts.

Genotyping by sequencing is one of the next-generation sequencing platforms that uses a simple, highly multiplexed system for constructing reduced representation libraries. It also uses inexpensive barcoding, reduces sample handling, requires fewer polymerase chain reaction and purification steps, and includes no size fractionation (Elshire et al., 2011). GBS can be applied to a wide array of organisms including plants for genome sequencing and SNP discovery, and is a rapid and inexpensive approach for trait mapping and association. With GBS, plant breeders can use techniques of molecular breeding by conducting genomic selection on any germplasm or species with or without prior knowledge of the genome in the species (Elshire et al., 2011; Sonah et al., 2013). The GBS platform is an advantageous approach for genome-wide SNP discovery, genetic map construction, linkage mapping, and genome-wide association in spinach.

Genetic diversity forms the raw material of plant breeding and is crucial for successful breeding programs (Jansen et al., 2006). Understanding the genetic diversity in one's crop allows a breeder to make informed choices when making crosses and when incorporating more variation into their program. Genetic diversity also plays an important role in association mapping by providing population structure information [kinship matrix (K-matrix)] to analyze loci association with traits (Khan, 2013; Khan and Korban, 2012).

Because the use of molecular markers in spinach has been limited up to this point, molecular diversity studies have also been limited (Hu et al., 2007). Kuwahara et al. (2012) analyzed 250 individuals from West Asia, East Asia, Japan, Europe, and the United States using SSR markers for six loci and found overall significant genetic differentiation among spinach from the different geographical regions. Diversity has also been observed among Iranian landraces of spinach, where high variation in morphological traits such as leaf shape, pedicle length, and the percentage of female plants were correlated to the variation in genotypes (Sabaghnia et al., 2014). Further contributions to the understanding of genetic diversity in spinach will be useful for spinach breeding efforts.

The objective of this study was to perform association analysis for bolting, plant height, and leaf erectness in the 288 accessions of USDA spinach collection.

Materials and Methods

Plant material and phenotyping. A total of 288 spinach accessions were used for the association analysis in this study (Supplemental Table 1). All seeds were kindly provided by David Brenner at the North Central Regional Plant Introduction Station, USDA-ARS, Iowa State University, Ames, IA, originally collected from 30 countries.

Phenotypic data of spinach bolting, plant height, and erectness of the 288 accessions were observed at the USDA-ARS research station in Salinas, CA, and can be downloaded from USDA-GRIN web site at https://npgsweb.ars-grin.gov/gringlobal/method.aspx?id=492382. For each accession, there were 10 plants grown in plastic pots (10 × 10 × 10 cm) with 2 sand: 1 soil (by volume) in a greenhouse. Plant height was measured as the height from soil/medium surface to the highest leaf tip of the plant 55 d after planting. For leaf erectness, leaves were rated "semiupright" if they were ≈45° from horizontal level and "upright" if they were closer to the upright position. An accession was deemed "early bolting" if any plant started stem elongation earlier than 60 d after planting, "intermediate" if bolting between 60 and 70 d, and "late bolting" after the 70 d.

Phenotypic data for plant height were analyzed using Microsoft (MS) Excel 2013 for the average, range, SD, SE, and coefficient of variation (CV). The CV, also known as relative SD, is a standardized measure of dispersion of a probability distribution or frequency distribution, where CV = SD/mean × 100. The distributions of bolting, plant height, and erectness were also drawn using MS Excel.

DNA extraction, GBS, and SNP discovery. Genomic DNA was extracted from fresh leaves of greenhouse-grown spinach plants using the hexadecyltrimethyl ammonium bromide method (Kisha et al., 1997). DNA sequencing was done by next-generation sequencing technologies using GBS (Elshire et al., 2011; Sonah et al., 2013). GBS was done using Illumina HiSEq 2000 at the Beijing Genome Institute (BGI), Hong Kong, China. Sequence assembly, mapping, and SNP discovery of GBS data were analyzed using SOAP family software (http://soap.genomics.org.cn/). The GBS data provided by BGI averaged 3.26 M short read and 283.74 Mbp data points for each spinach sample. The short reads of the GBS data were aligned to spinach genome reference Viroflay-1.0.1 (AYZV01) (http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV01#contigs) using SOAPaligner/soap2 (http://soap.genomics.org.cn/) and SOAPsnp v 1.05 was used for SNP calling (Li, 2011; Li et al., 2009). About one half-million SNPs were discovered from the GBS data among the 288 spinach germplasm accessions and the original SNP data were also provided by BGI. The SNP information was updated to spinach genome reference Spinach-1.0.3 (AYZV02) (http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV02) using BLAST after AYZV02 was released on 7 July 2015. The spinach accessions and SNPs were filtered before conducting genetic diversity and association analyses. If the spinach accession had greater than 35% missing SNP data, the genotype was removed from the panel. The SNP data were filtered by minor allele frequency >2%, missing data <25%, and heterozygous genotype <50%. After filtering, 1733 SNPs for 288 spinach accessions were used for genetic diversity and association analysis.

Population structure and genetic diversity. The model-based program STRUCTURE 2.3.4 (Pritchard et al., 2000) was used to assess the population structure of the 288 spinach accessions based on 1733 loci. To identify the number of populations (K), making up the structure of the data, the burn-in period was set at 10,000 with the Markov Chain Monte Carlo iterations and the run length set at 10,000 in an admixture model. The analysis then correlated allele frequencies independent for each run (Lv et al., 2012). Ten runs were performed for each simulated value of K, which ranged from 1 to 10. For each simulated K, the statistical value delta K was calculated using the formula described by Evanno et al. (2005). The optimal K was determined using STRUCTURE HARVESTER (Earl and von Holdt, 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. Each spinach accession was then assigned to a cluster (Q) based on the probability determined by the software that the genotype belonged in the cluster. The cutoff probability for assignment to a cluster was 0.50. Based on the optimum K, a Bar plot with "Sort by Q" was obtained to show the visual of the population structure among the 288 spinach accessions.

Genetic diversity was also assessed and the phylogenetic trees were drawn using MEGA 6 (Tamura et al., 2013) based on the maximum-likelihood (ML) tree method with the following parameters. Test of phylogeny: bootstrap method, number 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: 4, gaps/missing data treatment: use all sites, ML heuristic method: subtree-pruning-regrafting-extensive (level 5), initial tree for ML: make initial tree automatically (neighbor joining), and branch swap filter: moderate. To compare the results from the two software programs, during the drawing of the phylogeny trees by MEGA, the colored shape and branch of each spinach genotype were drawn using the same color, which was located at the cluster (Q) from STRUCTURE. For subtree 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 performed using TASSEL 5 software,

in which GLM, and MLM were used and compared (Bradbury et al., 2007; http://www.maizegenetics.net/tassel). GLM analysis incorporated population structure (Q-matrix) and MLM used both population structure (Q-matrix) and K-matrix in the association analysis (Bradbury et al., 2007; Shi et al., 2016). Q-matrix was estimated using STRUCTURE 2.3.4 (Pritchard et al., 2000) as described in above section in detail. Kinship (K-matrix) was estimated by the tool K-matrix built in Tassel 5 with Scald_IBS method. The QGene 4.3.10 was used to conduct SMR for all SNPs (Joehanes and Nelson 2008). Although QGene was developed for QTL mapping, it can also be used in association analysis through SMR. SMR for each SNP was estimated using QGene with 1733 SNP loci in 288 genotypes without Q and K matrices.

Results and Discussion

Phenotyping

Phenotypic data for bolting were classified as early, intermediate, or late. In this research, we only selected early and late-bolting types: 173 early and 115 late accessions were included (Supplemental Table 1; Fig. 1). Phenotypic data of plant height were measured in centimeters and they showed a near normal distribution (Supplemental Table 1; Fig. 2). The range of plant height was from 4.5 to 16.2 cm with a median of 8.6 cm and an average of 8.8 cm, and the SD of plant height was 1.9 with the SD error 0.0065. The CV was 21.3%, indicating there were significant genetic differences of plant height among the 288 spinach accessions (Supplemental Table 1). Phenotypic data for erectness were classified as SEMI or UP. Of the 288 accessions, 230 were SEMI and 58 were UP (Supplemental Table 1; Fig. 3).

Population structure

The population structure of the 288 spinach accessions was inferred using STRUCTURE 2.3.4 and the optimum K was K = 2 with the online tool STRUCTURE HARVESTER at http://taylor0.biology.ucla.edu/structureHarvester/ (Earl and von Holdt, 2012), as indicated by the highest delta K value (Fig. 4A). This indicated the presence of two main population clusters (Q1 and Q2) within the 288 spinach accessions. Figure 4B is the bar plot drawn to visualize the population structure where Q1 is red and Q2 is green. Each spinach accession was assigned to one of the two populations based on probabilities (P) given by STRUCTURE. Because some spinach accessions had similar *P* values between the two clusters, we defined the accession as Q1Q2 of admixture. There are 93 accessions in Q1 (32.3%), 129 accessions in Q2 (44.8%), and 66 accessions in the admixture Q1Q2 (22.9%) (Supplemental Table 1).

Genetic diversity was further analyzed using the 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 clusters and used the same colors as the population structure Q1 (red) and Q2 (green) from STRUCTURE 2.3.4 (Fig. 4B) to draw the subtrees of the phylogenetic tree in MEGA 6 plus the admixture Q1Q2 (Fig. 4C). The phylogenetic tree (Fig. 4C) from MEGA 6 was consistent with the structure populations (Q1 and Q2) from STRUCTURE 2.3.4, indicating that there were two well-differentiated genetic populations and admixture in the spinach panel plus the admixture Q1Q2 with the empty black square shape in the Fig. 4C.

To view the phylogenetic trees easily, we combined the spinach accession number, the accession original country, the accession geography region, and the structure population (cluster) into one taxon name for each spinach accession to draw the combined tree. For example, the taxon name, Ames23662_Afghanistan_Asia_Q1, includes the accession number Ames23662, which was originally collected from Afghanistan in Asia and assigned to cluster Q1. The combining taxon name for each spinach accession is shown in the Supplemental Table 1, and Supplemental Fig. 1. Because of the large size of the table and figures, they are listed in the supporting information. Viewing from Fig. 4 and Supplemental Fig. 1, the 288 spinach accessions showed a clear division when they were organized into two structured populations. Therefore, we used the Q matrix with two structures in the association mapping in TASSEL below.

In this manuscript, we used the STRUCTURE 2.3.4 (Pritchard et al., 2000) to determine population structure and pick up the k when the delta K value was highest. We also used MEGA 6 (Tamura et al., 2013) to analyze the genetic diversity and draw the phylogenetic trees for the same association panel, if both analyses from STRUCTURE and MEGA were matched; we assumed these were well-differentiated genetic populations and admixture in the panel. We then used Q-matrix with k vector in TASSEL for association analysis. STRUCTURE software has been a widely used program for association mapping in plants (Jin et al., 2010; Price et al., 2010; Pritchard et al., 2000; Shi et al., 2016; Upadhyaya et al., 2013; Zhu et al., 2008) and provides an effective correction for population stratification (Price et al., 2010). Population stratification is an issue

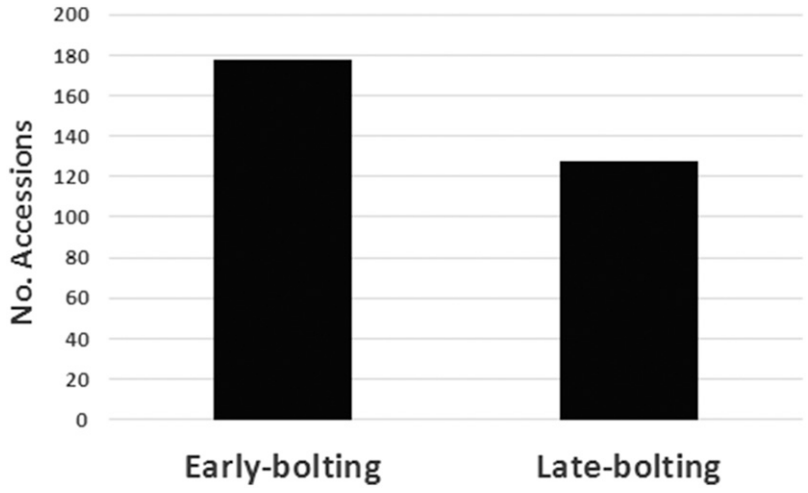

Fig. 1. The distribution of spinach bolting in 288 spinach accessions (an accession was deemed ''early bolting'' if any plant started stem elongation earlier than 60 d after planting, and ''late bolting'' after 70 d).

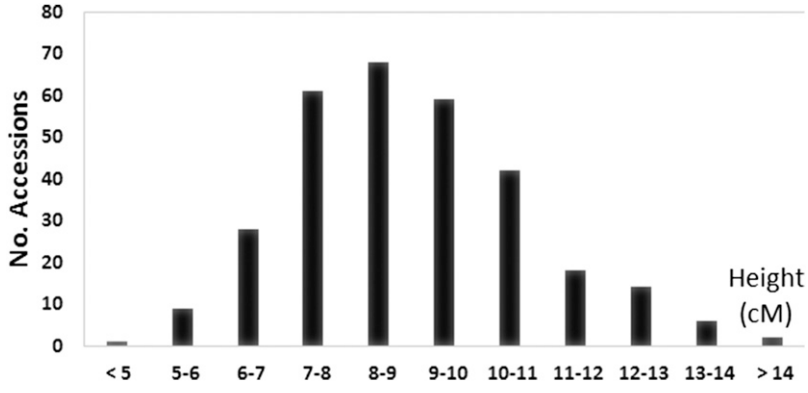

Fig. 2. The distribution of spinach plant height in 288 spinach accessions [plant height was measured as the height (cm) from ground to the highest leaf tip of the plant 55 d after planting].

that affects association mapping and many different methods and models of correcting for stratification have been developed (Freedman et al., 2004; Price et al., 2010; Pritchard et al., 2000). There are limited reviews on the impact of population stratification in association mapping (Freedman et al., 2004; Price et al., 2010). The Mixed Models by Price et al. (2010) is believed to be of future use in spinach association mapping.

Association analysis

SNP markers were identified for bolting, plant height, and erectness using three models, SMR, GLM, and MLM.

Bolting. SMR, GLM, and MLM approaches all identified three SNP markers, AYZV02001321_398, AYZV02041012_1060, and AYZV02118171_95, as having association with bolting with a P value <0.0001 (Table 1). The percentages of R^2 for the three SNP markers AYZV02001321_398, AYZV02041012_1060, and AYZV02118171_95 were 8.5%, 6.6%, and 6.6%, respectively, based on SMR. The GLM produced R^2 values of 8.7%, 6.8%, and 6.3%, respectively, and MLM was similar with 7.8%, 6.5%, and 6.1%, respectively. The smaller P value with not lower R^2 indicated that the three SNP markers were good markers, which may be validated for use in spinach breeding to select for late bolting through MAS.

Plant height. Eight SNP markers, AYZV02014270_540, AYZV02250508_2162, AYZV02091523_19842, AYZV02141794_376, AYZV02077023_64, AYZV02210662_2532, AYZV02153224_2197, and AYZV02003975_248, were associated with spinach plant height with P values <0.001 except AYZV02153224_2197 and AYZV02003975_248 based on MLM (Table 1). The percentages of R^2 ranged from 3.9% to 10.4% (Table 1). The SNP markers AYZV02014270_540 and AYZV02250508_2162 were excellent markers with P values <0.000001, <0.00001, and <0.0001 from SMR, GLM, and MLM, respectively. The R^2 was greater than 8.8%, 8.2%, and 7.0% from SMR, GLM, and MLM, respectively (Table 1), indicating the two SNP markers were strongly associated with spinach plant height and may be accurate markers for selection of plant height in spinach breeding through MAS after validation.

Erectness. Four SNP markers, AYZV02188832_229, AYZV02219088_79, AYZV02030116_256, and AYZV02129827_197, were associated with erectness (Table 1). SMR, GLM, and MLM did not show similar results for the four SNP markers. AYZV02188832_229 and AYZV02219088_79 were good markers with P value <0.001 except P = 0.00153 at MLM analysis for AYZV02219088_79 (Table 1). AYZV02030116_256 and AYZV02129827_197 showed association with P value < 0.006 except P = 0.01109 at SMR analysis for AYZV02129827_197. Therefore, these four SNP markers may provide a tool for selecting erectness in spinach molecular breeding.

In this study, the association studies were performed by using a compressed mixed linear model (Zhang et al., 2010) implemented in TASSEL 5 (Bradbury et al., 2007). The model incorporated population structure as fixed effects and cryptic relationship among individuals to define the variance structure of random individual genetic effects to control false positives. The analysis of population structure was conducted by using STRUCTURE software package to derive the Q matrix (Pritchard et al., 2000). The Qgene 4.3.10 software was used for SMR. Although QGene was developed for QTL mapping, it can also be used in association analysis through SMR. Although the different models SMR, GLM, and MLM did not provide the same results in our study, three SNP markers for bolting, eight markers for plant height, and four markers for erectness were found to be consistently associated with the traits. Currently, the available spinach genome reference Spinach-1.0.3 (AYZV02) (http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AYZV02) as released on 7 July 2015, represented about one-half of the spinach genome (Dohm et al., 2014; Minoche et al., 2015). A more comprehensive version of the spinach genome assembly may be made available publicly in 2016 (van Deynze, 2014; van Deynze et al., 2015; Allen van Deynze, personal communication), but unfortunately, the spinach whole genome sequences with physical maps are not available publicly as of this writing. After the whole genome sequences become publicly available, QTLs for bolting, plant height, or leaf erectness can be mapped to chromosome location.

Spinach bolting, plant height, and leaf erectness are important agronomic traits. Slow-bolting or late-bolting spinach can be used for year-round production because of its lack of sensitivity to photoperiod (Chun et al., 2000). And, slow bolting or late bolting is

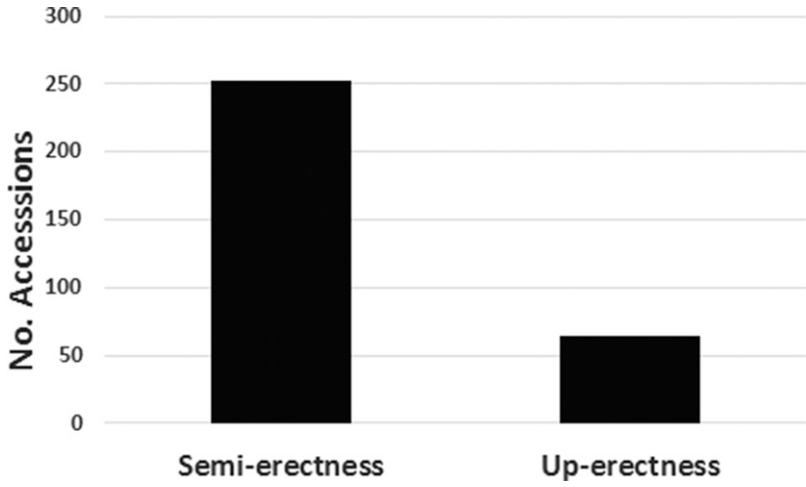

Fig. 3. The distribution of spinach erectness in 288 spinach accessions (for leaf erectness, leaves were rated "semiupright" if they were ≈45° from horizontal level and "upright" if they were closer to the upright position).

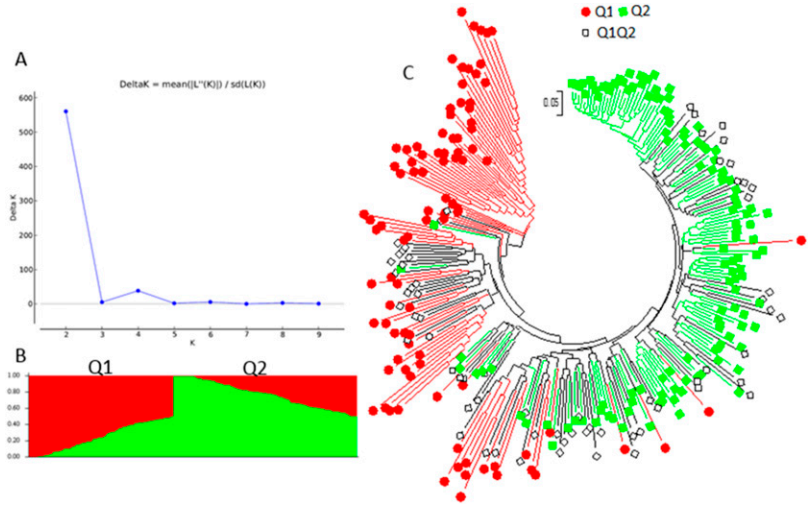

Fig. 4. Model-based populations in the association panel (**A**) Delta K values for different numbers of populations assumed (K) in the STRUCTURE analysis. (**B**) Classification of spinach accessions into two populations using STRUCTURE 2.3.4. The distribution of the accessions to different populations is indicated by the color code (Q1: red and Q2: green). (**C**) Maximum likelihood tree of the 288 accessions drawn by MEGA 6. The color codes for each population are consistent in **B** and **C**, and the empty black square as the admixture Q1Q2.

Table 1. Fifteen single nucleotide polymorphism (SNP) markers associated with bolting, plant height, and leaf erectness identified from three modes using QGene4 and Tassel in 288 spinach accessions.

Spinach genome Spinach-1.0.3 information				Viroflay-1.0.1		*P* value			R^2 (%)			Trait
SNP name[z]	SNP type	Contig at AYZV02 project	SNP position	Contig at AYZV01 project	SNP position	SMR[y]	GLM[y]	MLM[y]	SMR	GLM	MLM	
AYZV02001321_398	C/A	AYZV02001321	398	AYZV01001038	398	1.39E-06	1.13E-06	1.05E-05	8.5	8.7	7.8	Bolting
AYZV02041012_1060	G/A	AYZV02041012	1,060	AYZV01031624	1,060	5.28E-05	4.04E-05	9.93E-05	6.6	6.8	6.5	
AYZV02118171_95	G/A	AYZV02118171	95	AYZV01088923	95	1.75E-04	2.93E-04	5.15E-04	6.6	6.3	6.1	
AYZV02014270_540	A/G	AYZV02014270	540	AYZV01011130	540	9.87E-07	2.62E-06	5.32E-05	8.9	8.3	7.1	Plant height
AYZV02250508_2162	T/A	AYZV02250508	2,162	AYZV01180397	2,162	1.23E-07	6.53E-07	7.24E-05	10.4	9.2	7.1	
AYZV02091523_19842	A/G	AYZV02091523	19,842	AYZV01069590	19,842	5.76E-06	1.21E-05	3.65E-04	8	7.5	6.1	
AYZV02141794_376	A/G	AYZV02141794	376	AYZV01105690	376	1.61E-04	2.05E-04	4.00E-04	6.4	6.1	6.1	
AYZV02077023_64	G/T	AYZV02077023	64	AYZV01058838	64	3.84E-04	3.93E-04	4.67E-04	6.3	6.2	6.3	
AYZV02210662_2532	A/G	AYZV02210662	2,532	AYZV01152613	2,532	4.46E-05	1.36E-04	5.56E-04	6.6	5.8	5.2	
AYZV02153224_2197	T/C	AYZV02153224	2,197	AYZV01113619	2,197	3.74E-04	4.93E-04	0.00157	4.4	4.2	3.9	
AYZV02003975_248	T/A	AYZV02003975	248	AYZV01003134	248	2.61E-04	5.63E-04	0.00165	6	5.4	4.9	
AYZV02188832_229	G/T	AYZV02188832	229	AYZV01137843	229	2.34E-04	0.00107	7.15E-04	6.1	5	6.7	Erectness
AYZV02219088_79	C/A	AYZV02219088	79	AYZV01158294	79	1.72E-04	1.63E-04	0.00153	6.1	6.1	5.2	
AYZV02030116_256	T/C	AYZV02030116	256	AYZV01023368	256	0.00125	5.25E-04	0.00202	4.9	5.5	5.5	
AYZV02129827_197	A/C	AYZV02129827	197	AYZV01097131	197	0.01109	3.87E-05	0.00678	3	6.6	3.6	

[z]SNP name is defined as the contig name plus the SNP position on the contig.

[y]SMR = single marker regression using the QGene 4.3.10 (Joehanes and Nelson, 2008); GLM = regression linear model and MLM = mixed linear model using TASSEL 5 (Bradbury et al., 2007; http://www.maizegenetics.net/tassel).

also good for commercially grown spinach which is often cut multiple times during the growing season (Morelock and Correll, 2008) and increases the overall yield. Commercial spinach cultivation is highly mechanized (Morelock and Correll, 2008; Koike et al., 2011), and spinach cultivars with taller plant height and erect leaves are generally preferred to accommodate high-density spinach production and mechanical harvesting. From this research, eight accessions, PI103063, PI169678, PI169684, PI171863, PI171865, PI174386, PI175929, and PI648963, were identified as late bolting and had erect leaves with 9-cm plant height and they are good sources as parents in spinach breeding program.

Conclusions

Three SNP markers, AYZV02001321_398, AYZV02041012_1060, and AYZV02118171_95, were identified to be associated with bolting. Eight SNP markers, AYZV02014270_540, AYZV02250508_2162, AYZV02091523_19842, AYZV02141794_376, AYZV02077023_64, AYZV02210662_2532, AYZV02153224_2197, and AYZV02003975_248, were found to be associated with plant height. Four SNP markers, AYZV02188832_229, AYZV02219088_79, AYZV02030116_256, and AYZV02129827_197, were associated with erectness. These SNP markers may provide a tool to be used in spinach molecular breeding to select for bolting, plant height, and erectness through MAS.

Literature Cited

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635.

Chan-Navarrete, R., O. Dolstra, M. van Kaauwen, E.T.L. van Bueren, and C.G. van der Linden. 2016. Genetic map construction and QTL analysis of nitrogen use efficiency in spinach (*Spinacia oleracea* L.). Euphytica 208:621–636.

Chun, C., A. Watanabe, H. Kim, T. Kozai, and J. Fuse. 2000. Bolting and growth of *Spinacia oleracea* L. can be altered by modifying the photoperiod during transplant production. HortScience 35:624–626.

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

Earl, D.A. and B.M. von Holdt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resources 4:359–361.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

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

Freedman, M.L., D. Reich, K.L. Penney, G.J. McDonald, A.A. Mignault, N. Patterson, S.B. Gabriel, E.J. Topol, J.W. Smoller, C.N. Pato, M.T. Pato, T.L. Petryshen, L.N. Kolone, E.S. Lander, P. Sklar, B. Henderson, J.N. Hirschhorn, and D. Altshuler. 2004. Assessing the impact of population stratification on genetic association studies. Natl. Genet. 36:388–393.

Goreta, S. and D.I. Leskovar. 2006. Screening spinach cultivars for white rust and bolting. HortTechnology 16:162–166.

Hu, J., B. Mou, and B.A. Vick. 2007. Genetic diversity of 38 spinach (*Spinacia oleracea* L.) germplasm accessions and 10 commercial hybrids assessed by TRAP markers. Genet. Resources Crop Evol. 54:1667–1674.

Jansen, J., H. Verbakel, J. Peleman, and T.J. van Hintum. 2006. A note on the measurement of genetic diversity within genebank accessions of lettuce (*Lactuca sativa* L.) using AFLP markers. Theor. Appl. Genet. 112:554–561.

Jin, L., Y. Lu, P. Xiao, M. Sun, H. Corke, and J. Bao. 2010. Genetic diversity and population structure of a diverse set of rice germplasm for association mapping. Theor. Appl. Genet. 121:475–487.

Joehanes, R. and J.C. Nelson. 2008. QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics 24:2788–2789.

Khan, M.A. 2013. Association mapping using TASSEL software. 31 Dec. 2015. <http://articles.extension.org/pages/62755/association-mapping-using-tassel-software>.

Khan, M.A. and S. Korban. 2012. Association mapping in forest trees and fruit crops. J. Expt. Bot. 63:4045–4060.

Khattak, J., A.M. Torp, and S.B. Andersen. 2006. A genetic linkage map of *Spinacia oleracea* and localization of a sex determination locus. Euphytica 148:311–318.

Kisha, T., C.H. Sneller, and B.W. Diers. 1997. Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci. 37:1317–1325.

Koike, S., M. Cahn, M. Cantwell, S. Fennimore, M. Lestrange, E. Natwick, R. Smith, and E. Takele. 2011. Spinach production in California. Univ. California Agri. Natural Resources. Publ. 7212.

Kumar, S., T.W. Banks, and S. Cloutier. 2012. SNP discovery through next-generation sequencing and its applications. Intl. J. Plant Genomics, Article ID 831460 <http://dx.doi.org/10.1155/2012/831460>.

Kuwahara, K., R. Suzuki, Y. Ito, T. Mikami, and Y. Onodera. 2012. An analysis of genetic differentiation and geographical variation of spinach germplasm using SSR markers. Plant Genet. Resour. 12:185–195.

Lakew, B., R. Henry, S. Ceccarelli, S. Grando, J. Eglinton, and M. Baum. 2013. Genetic analysis and phenotypic associations for drought tolerance in *Hordeum spontaneum* introgression lines using SSR and SNP markers. Euphytica 189:9–29.

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, R., C. Yu, Y. Li, T. Lam, S. Yiu, K. Kristiansen, and J. Wnag. 2009. SOAP2: An improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967.

Lucas, M., J. Ehlers, B. Huynh, N. Diop, P. Roberts, and T. Close. 2013. Markers for breeding heat-tolerant cowpea. Mol. Breed. 31:529–536.

Lv, J., J. Qi, Q. Shi, D. Shen, S. Zhang, A. Zhang, G. Shao, H. Li, Z. Sun, Y. Weng, Y. Shang, X. Gu, X. Li, X. Zhu, J. Zhang, R. van Treuren, W. van Dooijeweert, Z. Zhang, and S. Huang. 2012. Genetic diversity and population Structure of cucumber (*Cucumis sativus* L). PLoS One 7:46919.

Minoche, A.E., J.C. Dohm, J. Schneider, D. Holtgrawe, P. Viehover, M. Montfort, T.R. Sorensen, B. Weisshaar, and H. Himmelbauer. 2015. Exploiting single-molecule transcript sequencing for eukaryotic gene prediction. Genome Biol. 16:184.

Morelock, T.E. and J.C. Correll. 2008. Spinach, p. 189–218. In: J. Prohens and F. Nuez, (eds.). Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae. Springer, New York, NY.

Nemli, S., T. Asciogul, H. Kaya, A. Kahraman, D. Esiyok, and B. Tanyolac. 2014. Association mapping of five agronomic traits in the common bean (*Phaseolus vulgaris* L.). J. Sci. Food Agr. 94:3141–3151.

Price, A.L., N.A. Zaitlen, D. Reich, and N. Patterson. 2010. New approaches to population stratification in genome-wide association studies. Nat. Rev. Genet. 11:459–463.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Sabaghnia, N., H. Asadi-Gharneh, and M. Janmohammadi. 2014. Genetic diversity of spinach (*Spinacia oleracea* L.) landraces collected in Iran using morphological traits. Acta Agr. Slov. 103:101–111.

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

Sonah, H., M. Bastien, E. Iquira, A. Tardivel, G. Legare, B. Boyle, É. Normandeau, J. Laroche, S. Larose, M. Jean, and F. Belzile. 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PLoS One 8:e54603.

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

Thomson, M., A. Ismail, S. McCouch, and D. Mackill. 2010. Marker assisted breeding, p. 451–469. Abiotic stress adaptation in plants: Physiological, molecular, and genetic foundation. Springer, The Netherlands.

Upadhyaya, H., Y. Wang, R. Sharma, and S. Sharma. 2013. Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. Theor. Appl. Genet. 126:1649–1657.

van Deynze, A. 2014. A De Novo draft assembly of spinach using pacific biosciences technology. Plant & animal genomics XXII conference, 10–15 Jan. 2014, San Diego, CA. 21 Mar. 2016. <http://aa314.gondor.co/webinar/a-de-novo-draft-assembly-of-spinach-using-pacific-biosciences-technology/>.

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

Zhang, Z., E. Ersoz, C.Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, P.J. Bradbury, J. Yu, D.K. Arnett, J.M. Ordovas, and E.S. Buckler. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. 42:355–360.

Zhu, C., M. Gore, E. Buckler, and J. Yu. 2008. Status and prospects of association mapping in plants. Plant Genome 1:5–20.