Genome-Wide Association Studies for Spot Blotch (*Cochliobolus sativus*) Resistance in Bread Wheat Using Genotyping-by-Sequencing

Muhammad Jamil,[†] Aamir Ali, Alvina Gul, Abdul Ghafoor, Amir M. H. Ibrahim, and Abdul Mujeeb-Kazi

First and second authors: Department of Botany, University of Sargodha, Sargodha, Pakistan; third author: Atta-ur-Rehman School of Applied Biosciences, National University of Science and Technology, Islamabad, Pakistan and United States Department of Agriculture–Agricultural Research Service, Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66506, USA; fourth author: Plant Genetic Resources Institute, National Agriculture Research Center, Islamabad, Pakistan; fifth author: Soil and Crop Sciences Department, Texas A&M University, TX 77843-2474; and sixth author: Texas A&M University, Amarillo, TX 79106.

Accepted for publication 26 May 2018.

ABSTRACT

Spot blotch is a severe biotic menace of wheat caused by *Cochliobolus sativus* (syn. *Bipolaris sorokiniana*). Spot blotch is liable to major yield losses in warm humid regions. A genome-wide association study using genotyping-by-sequencing (GBS) markers was conducted to identify genomic regions associated with spot blotch resistance in a diversity panel of 159 spring wheat genotypes. In total, 87,096 GBS markers covering the whole genome, with an average polymorphism information content value of 0.276, were applied. Linkage disequilibrium (LD) analysis indicated that the LD decay extent was approximately 100 Mbp. The panel was evaluated

for disease severity (DS) and area under disease progress curve (AUDPC) for 2 years. In total, 24 marker-trait associations (MTA) were identified for DS and AUDPC of spot blotch, with 11 on chromosome 5B, 3 on 3A, 2 on 6B, and 1 each on 1A, 2A, 1D, 2D, 4B, 5A, 7A, and 7B. A marker on chromosome 7B significantly explained 14% of the phenotypic variation of spot blotch severity as well as 11% of AUDPC. Five markers—three on chromosome 5B, one on 3A, and one on 7B—were associated with both DS and AUDPC with R^2 ranging from 8 to 12%. Significant MTA can be utilized to develop wheat germplasm with resistance to spot blotch.

Spot blotch (SB), caused by *Cochliobolus sativus* (S. Ito & Kurib.) Drechsler ex Dastur, is a major threat to wheat production in warm, humid regions. Previous studies showed that SB is responsible for up to 50% yield reduction under disease-prone conditions, which is a threat to global food security (Villareal et al. 1995). Now this disease is heading toward other major wheat-growing areas worldwide, including Eastern India, Southeast Asia, Latin America, China, and Africa (Kumar et al. 2010). Selecting for host plant resistance is the most effective measure to control the disease. Resistance can be transferred in a qualitative as well as in a quantitative manner (Kumar et al. 2009, 2010). Resistance genes can lose their effectiveness with the passage of time due to genetic mutation in the pathogen. These threats necessitate the identification of new resistance genes. Further use of marker-assisted selection will improve selection accuracy of disease resistance in wheat (Gurung et al. 2014).

Previously, several studies have focused on identifying genomic regions associated with SB resistance through biparental mapping of quantitative trait loci (QTL) (Kumar et al. 2009; Lillemo et al. 2013) and genome-wide association study (GWAS) (Wang et al. 2017). QTL for SB resistance on chromosomes 2A, 2B, 5B, 6D, and 7D had been reported earlier by Kumar et al. (2009, 2010). Lillemo et al. (2013) reported colocalization of QTL with *Lr34* on chromosome 7D and *Lr46*. Recently, QTL on chromosomes 1B, 3B, and 5A have been reported using simple-sequence repeat markers (Zhu et al. 2014). Diversity Arrays Technology (DArT) was also used to identify the loci associated with SB resistance (Adhikari et al. 2012). More recently, Gurung et al. (2014) identified QTL

on chromosomes 1B, 5A, 5B, 6B, and 7B for SB resistance using high-density single-nucleotide polymorphism (SNP) arrays.

However, all previously mentioned genotyping approaches each provide a limited number of markers that may not be able to cover an entire genome at the desired depth. Development of next-generation sequencing (NGS) technologies has enabled researchers to discover an increased number of SNP that will provide a better understanding of the genetic structure of complex traits for crop improvement (Atwell et al. 2010). NGS has provided an effective tool to develop high-throughput and high-density genotyping platforms for plant species (Rasheed et al. 2017). Genotyping-by-sequencing (GBS) has been highlighted as a novel genotyping platform for association studies in species such as wheat with complex genomes (Poland et al. 2012a). GBS permits high-quality, population-specific SNP extracted from ascertainment bias, which remains a big problem in fixed SNP arrays (Bajgain et al. 2016). In GBS, complexity of large genomes can be reduced using restriction digestion of enzymes and only a selected subset of informative sequences in the genome is sequenced using NGS for SNP calling (Elshire et al. 2011; Poland et al. 2012b). Using this technology, Saintenac et al. (2013) constructed a comprehensive sequence-based wheat map and integrated GBS-SNP with other markers, which is a useful resource for QTL mapping and high-throughput marker development. The objective of the present study was to conduct GWAS using GBS methodology to identify chromosome regions associated with SB resistance in spring wheat from the International Maize and Wheat Improvement Center (CIMMYT).

MATERIALS AND METHODS

Plant materials and SB phenotyping. In total, 159 spring bread wheat accessions from CIMMYT were used (Supplementary Table S1). All the materials were evaluated in two field experiments (2013–14 and 2014–15) at the National Agriculture Research Center (NARC), Islamabad, Pakistan. The percentage of

[†]Corresponding author: E-mail: jshahid80@yahoo.com

*The e-Xtra logo stands for “electronic extra” and indicates that two supplementary figures and three supplementary tables are published online.

SB severity was recorded at three different plant growth stages: GS63, -69, and -77 on Zadoks' scale (Zadoks et al. 1974). Area under the disease progressive curve (AUDPC) was also calculated to reflect the disease progress over time. Each accession was planted in a 2-m row with 5 cm between plants in each row and 30 cm between rows. Experiments followed a randomized complete block design with three replicates. Resistant cultivar Chirya 3 and susceptible cultivar Ciano 79 were used as check cultivars. Ciano 79 was also planted after every 20th row and along the border of the field to increase inoculum density.

The most virulent aggressive isolates of *Bipolaris sorokiniana* from the Crop Disease Research Institute (CDRI) at NARC, Islamabad were used as inoculum for the experiments to create artificial epiphytotic conditions. The most virulent pure culture of the aggressive local isolate of *B. sorokiniana* identified by the CDRI at NARC, Islamabad was used in the study (Asad et al. 2007, 2009). This inoculum was multiplied on wheat grains and spores were collected in water. The disease was induced by a uniform spray at three different GS (tillering: GS20, flag leaf emergence: GS37, and anthesis: GS65 on Zadoks' scale) on the whole plots during the late noon with a spore water suspension set to 10⁴ spores/ml of water (Kumar et al. 2009). Further agronomic practices were followed as performed by Kumar et al. (2009). Disease severity was estimated using a double-digit rating scale (00 to 99), as described by Mujeeb-Kazi et al. (2007). Percent severity was calculated as Percentage of severity = D1/9 × D2/9 × 100, where digit 1 (D1) is infection along the height of the plant (0 = no disease, 1 = infection is at the bottom near the ground, and 9 = infection up to the flag leaf) and digit 2 (D2) is the severity of the infection estimated across all leaves of all the plants up to the level indicated by the first digit (0 = no infection, 1 = 10% area infected, and 9 = ≥90% leaf area of all the leaves in all the plants infected).

TABLE 1. Frequency of the markers, polymorphism information contents (PIC), and genetic diversity genotyping-by-sequencing markers

Chromosomes	Markers	PIC ^a	Diversity
Genome A			
1	3,756	0.283	0.337
2	4,731	0.283	0.339
3	4,187	0.268	0.319
4	3,961	0.275	0.327
5	3,849	0.271	0.321
6	3,614	0.284	0.339
7	5,436	0.265	0.314
Subtotal	29,534 ^b	0.276 ^c	0.328 ^c
Genome B			
1	4,672	0.275	0.327
2	6,232	0.285	0.342
3	5,656	0.278	0.330
4	3,326	0.276	0.325
5	5,215	0.280	0.333
6	5,131	0.278	0.333
7	5,484	0.275	0.327
Subtotal	35,716	0.278	0.331
Genome D			
1	2,649	0.272	0.316
2	3,879	0.271	0.314
3	3,344	0.272	0.312
4	1,990	0.279	0.321
5	3,042	0.275	0.318
6	2,817	0.280	0.328
7	4,125	0.261	0.303
Subtotal	21,846	0.273	0.316
Total	87,096	0.276 ^d	0.325 ^d

^a Range: ≥PIC 0.5 = 4,406 markers, ≥0.6 PIC = 1,208 markers, and ≥0.7 PIC = 172 markers.
^b Total markers.
^c Mean value of subgenome.
^d Mean value across the genome A, B, and D.

Fig. 1. Simple linear regression of severity against area under disease progress curve (AUDPC) in A, 2014 and B, 2015; box plots for spot blotch disease severity and (AUDPC) in years C, 2014 and D, 2015; and frequency distribution of 159 accessions into H = highly resistant, R = resistant, MR = moderately resistant, MS = moderately susceptible, S = susceptible, and HS = highly susceptible for E, severity and F, AUDPC.

Fig. 2. Single-nucleotide polymorphisms identified with 87,096 genotyping-by-sequencing markers.

Severity displayed in all the leaves of a row was recorded at three different stages: GS63 (beginning of anthesis to half complete), GS69 (anthesis complete), and GS77 (late milking stage) on the Zadoks et al. (1974) scale. AUDPC was calculated from the disease severity (DS) data collected at GS63, GS69, and GS77 using the following formula (Roelfs 1992): $AUDPC = \sum_{i=1}^n \left[\frac{(Y_i + Y_{i+1})}{2} \right] \times (t_{i+1} - t_i)$, where Y_i = disease level at time t_i , $(t_{i+1} - t_i)$ = time (days) between two disease scores, and n = number of days on which SB was recorded. Genotypes with $\leq 42.5\%$ DS and $\leq 1,349.11$ AUDPC were declared to be resistant genotypes.

GBS-SNP genotyping. All the accessions were genotyped using GBS. DNA was extracted using a kit from Prima Scientific (Bangkok, Thailand). GBS was conducted following a protocol by Mascher et al. (2013), with some modification such as E-gel size selection, and sequencing on the Ion Torrent Proton System using Ion PI v3 chips and Ion PI Hi-Q Chef Kit with 80 pM library input.

Fig. 3. A, Population structure of 159 genotypes indicated with plot of ΔK as second-order rate of change, highlighting ΔK value = 6 as compared with assumed $K = 15$; B, six clusters loaded on principal component axis; and C, bar plot indicating membership coefficient (Q value at Y-axis), where each vertical line is one genotype highlighting proportion from each of the six clusters. The appropriate cluster has a high proportion of the respective shade. Cluster 1 (Q1) = uppermost medium gray, cluster 2 (Q2) = second from top medium gray, cluster 3 (Q3) = black, cluster 4 (Q4) = lower medium gray, cluster 5 (Q5) = white, and cluster 6 (Q6) = pale gray.

SNP calling was performed using the TASSEL 5 pipeline by aligning reads against a Chinese Spring (CS) whole wheat genome assembly (WGA 0.4; International Wheat Genome Sequencing Consortium). SNP were called with minimum allele frequency set at 0.5 and minimum locus coverage set at 0.2. Missing data were imputed using the K-nearest-neighbors imputation approach method in TASSEL 5 (Money et al. 2015).

Population structure analysis. The genotypic data were subjected to numerical transformation with XLSTAT 2010 and then converted to the required format of the STRUCTURE 2.3.4 software (Pritchard et al. 2000). The admixture model was adjusted with burn-in period length of 100,000 followed by 500,000 Markov Chain Monte Carlo (MCMC) replications. The subpopulation test range was kept from K1 to K15, each with 10 iterations (runs) (Gurung et al. 2014). The ΔK approach was used to assess the actual subpopulations (Earl 2012). Average logarithm of the probability of observed likelihood [LnP(D)] was calculated along with the

standard deviation from the output summary. LnP(D) for each step of the MCMC was calculated for each class (K = 1 to 15) by computing the log likelihood of the data. In total, 150 LnP(D) values were distributed according to multiple runs (10 runs for each K). To avoid significant overlapping, ΔK was confirmed by the Evanno et al. (2005) method using the STRUCTURE HARVESTER program (Earl 2012).

Linkage disequilibrium. A pairwise linkage disequilibrium (LD) analysis was performed on the filtered set of SNP from 87,096 markers. Including minor SNP states, 1,386 loci were captured by keeping a minimum count of 152 of 159 genotypes, keeping allelic frequencies ranging from 0.065 to 1.0. For analysis of the full matrix of 1,386 loci, 959,112 comparisons were made by treating heterozygous calls as missing. LD plots were generated for each chromosome separately using TASSEL software. For each marker pair, R^2 was calculated and the base pair distance below 200 Mb was further magnified.

Fig. 4. Dendrogram of 159 bread wheat genotypes using the neighbor-joining method.

Marker-trait association. Marker-trait association (MTA) was performed for SB severity and AUDPC using a mixed-linear model (MLM). For MLM, a Q+K model that considers both kinship (K-matrix) and the population structure (Q-matrix) was used to determine the population structure using STRUCTURE 2.3.4 software. Averaged phenotypic data of 2 years, covariates (Q-matrix), and genotypic data were merged through intersection join in TASSEL. Merged files and kinship matrix were coselected to run MLM. Manhattan plots were generated by P value = 3 ($-\log_{10}$) as a threshold.

RESULTS

Evaluation of SB resistance of the diversity panel. The coefficients of determination between final DS and AUDPC were high ($R^2 = 89$ and 83%) for the years 2014 and 2015, respectively (Fig. 1A and B). On average, the SB severity of 40 to 42% and AUDPC of 1,216 to 1,491 were recorded for 2 years (Fig. 1C and D). Final DS and AUDPC of tested accessions in the panel showed near-normal distributions (Fig. 1E and F). Among 159 accessions, 21 and 23 were resistant, with final DS cut-off $< 29\%$ and AUDPC < 937 , respectively. On the basis of severity ($>55.5\%$) and AUDPC ($>1,760$), 15 and 24 genotypes, respectively, were sorted as susceptible (Fig. 1E and F). Further detailed descriptive statistics along with disease score of resistant and susceptible check cultivars has been given in Supplementary Table S2. Detailed temperature and rainfall information during the two test seasons is shown in Supplementary Figure S1.

Analysis of variance revealed that genotypic effect and genotype-year interaction were significant for both final DS and AUDPC (Supplementary Table S3). The heritability was calculated for severity and AUDPC were observed to be moderate (i.e., 0.52 and 0.53, respectively).

SNP distribution in the wheat genome. In total, 87,096 SNP were identified in the diversity panel. The B genome had the most SNP (41%), followed by the A genome (34%) and the D genome (25%) (Table 1). Among 21 chromosomes, chromosome 2B had the highest marker density (6,232 markers), followed by 3B (5,656 markers). Polymorphism information content (PIC) value was calculated for each marker, with a mean PIC of 0.28 for all SNP analyzed. A similar PIC was observed for three genomes. The average gene diversity was 0.32. Among 20 types of SNP identified (Fig. 2), SNP G/A was most frequent (12.2%) and SNP C/T was second-most frequent.

Population structure and LD. Population structure analysis divided the panel into six subpopulations (Fig. 3A and C) and principal component analysis also grouped 159 accessions into six classes (Fig. 3B), which were confirmed by cluster analysis using the neighbor-joining method (Fig. 4). Group I included 14 accessions, including the susceptible check Ciano 79. Group II included 50 accessions, group III included 12 accessions, group IV included 14 accessions, group V included 44 accessions, and group VI included 25 accessions (Fig. 4). The maximum dissimilarity was observed between the accession Ciano 79 and accession H.Sat-19.

Whole-genome pairwise LD analysis identified a total of 663,978 SNP pairs, and 43,391 of them showed significant LD (Table 2).

Among them, 344,444 (51.87%) SNP pairs were identified within the distance of 200 Mb, 12.08% (41,617) of them showed significant LD (P value ≤ 0.001), and 57.45% of them showed complete LD ($r^2 = 1$). The mean r^2 for whole-genome LD was 0.045. If $r^2 > 0.2$ is used as threshold value, as suggested by Breseghello and Sorrells (2006), 15.61% of pairs were found to be above this threshold value (Fig. 5). Smoothing was performed with a locally weighted least-square curve which showed decay at the distance of approximately 80 Mbp. Distance from 0 to 200 Mbp was further reanalyzed and showed the decay at approximately 30 Mbp.

MTA for SB resistance. In total, 24 MTA were identified, with 7 for SB severity, 12 for AUDPC, and 5 for both traits. As a whole, these MTA covered 16 chromosomal regions (Supplementary Fig. S2). On chromosome 5B, six GBS markers were found associated

Fig. 5. Linkage disequilibrium (LD) scatter plot (pairwise), showing squared allele-frequency correlation (r^2) decay along the base pair distance between pairs of genotyping-by-sequencing markers. The locally weighted polynomial regression-based representing decay of r^2 along physical distance is presented. LD critical thresholds ($r^2 = 0.2$ and $r^2 = 0.4$) are indicated by two horizontal lines. Megabase pair distance (Mbp) are shown on the x-axis and coefficient of determination (R^2) on y-axis.

TABLE 2. Summary of linkage disequilibrium (LD) statistics in 159 bread wheat genotypes using genotyping-by-sequencing markers

Distance (Mbp) ^a	Total pairs	Significant pairs	Significant (%)	Mean r^2	Pairs in complete LD	Pairs (%) in LD > 0.2	Mean of $r^2 > 0.2$	Pairs in LD > 0.2
0 to 200 Mbp	344,444	41,617	12.08	0.10	23,911	12.23	0.75	42,132
200 < 400 Mbp	117,117	820	0.70	0.03	0	2.38	0.28	2,795
400 < 600 Mbp	141,145	860	0.61	0.03	0	0.86	0.29	1,218
600 Mbp <	61,272	94	0.15	0.02	0	0.13	0.32	81
Total pairs	663,978	43,391	13.55	0.045	23,911	15.61	0.4	46,226

^a Mbp = million base pairs.

with AUDPC, two for severity, and three for both traits. Three GBS markers conferring AUDPC were identified on chromosome 3A, unraveling 13, 8, and 10% AUDPC variation, respectively (Table 3). Another genomic region on chromosome 7B conferred both traits. On chromosome 6B, two markers were identified for SB resistance, explaining 12 and 8% variability. One marker was found associated with severity on each of the chromosomes 1A, 2A, and 5A, showing 10, 10, and 11% R^2 , respectively. A marker on chromosome 3A (S3A_738506623) was identified with 8% R^2 for each AUDPC and final DS.

DISCUSSION

Phenotypic evaluation for SB resistance. The diversity panel included well-reported resistant check cultivar Chirya 3 (Kumar et al. 2010; Singh et al. 2015), released in 1992, and susceptible check cultivar Ciano 79, released in 1979. The diversity panel also included Sonalika and Francolin#1, which have been reported previously as susceptible and moderately susceptible, respectively (Adhikari et al. 2012; Kumar et al. 2009; Sharma et al. 2006; Singh et al. 2015). Francolin#1 is a CIMMYT cultivar that is known to be leaf rust and stripe rust resistant (Lan et al. 2014) but was used by Singh et al. (2015) as a moderately susceptible check. SB severity ranged from 1.23 to 89% and was well distributed across all germplasm. In previous studies, evaluation was performed in biparental lines such as recombinant inbred lines (Kumar et al. 2009) and backcross derivatives (Vasistha et al. 2016), and the severity score ranged from 20 to 88%. Neupane et al. (2007) produced a biparental population using Chirya 3 as a resistant parent and a disease score up to only 10% was observed. Although the genotype-year interaction was significant, it is important to keep in mind that wheat genotypes become more vulnerable to SB severity if temperature rises to 26°C at late milking stage (Kumar et al. 2009). Percentage of disease severity remained a primary indicator

for disease assessment (Villareal et al. 1995). Disease severity progress with time can be summed up in the form of AUDPC that is the elaborated measurement and precise expression of SB phenotypic evaluation (Jeger 2004). In our study, higher mean values for AUDPC were observed as compared with the previous study (Kumar et al. 2009; Kumar et al. 2015) performed on 341 single-seed descendants. This might be due to the multiple parentages of the germplasm panel used. The heritability for SB severity and AUDPC was 0.47 and 0.49, respectively, which was quite a bit higher than the previous studies performed by Sharma et al. (1997). This indicates that influence of environment is highly significant, which might be due to temperature differences at the time of disease development.

Genome coverage, population structure, and LD. The increase in marker density provided more opportunity to precisely identify genomic regions associated with targeted traits. The six groups observed during phylogenetic investigation were significantly different in their phenotypic and genotypic makeup. This shows that the population under study had enough diversity to identify multiple alleles due to the higher number of resistance sources. The susceptible check cultivar Ciano 79 along with another well-known susceptible check, Sonalika (Kumar et al. 2009; Singh et al. 2015), was observed in group I. The resistant check cultivar Chirya 3 was placed in group II and had the largest number of genotypes, mostly with WBL1, ROLF, PASTOR, PBW, and FRET in their parentage, providing the clues for the resistant genetic resources. Wheat accessions Kukuna, Milan, Kauz, and Kachu were the major contributors in most of the genotypes constituting group III and group IV. The moderately resistant check cultivar Francolin #1 (Singh et al. 2015) was observed in group V, along with other genotypes with Attila and Vorb in their parentage.

Synthetic wheat is a milestone in expanding the genetic pool to widen the diversity by introducing D-genome diversity (Ogbonnaya et al. 2013) through introgression of *Aegilops tauschii*, also known

TABLE 3. Marker-trait associations for severity and area under the disease progress curve (AUDPC) using mixed-linear model (Q+K model) approach in 159 diverse bread wheat lines^a

Trait, marker	Chr	Position	SNP	MAF	P (–log10)	Marker R^2	Known gene (LD: R^2)
AUDPC							
S1D_479711997	1D	479711997	A/G	0.18	4.59	0.11	...
S2D_389463371	2D	389463371	G/A	0.11	3.19	0.10	<i>Vrn-A1</i> (0.041)
S3A_180419285	3A	180419285	A/G	0.21	3.19	0.13	<i>Vrn-A1</i> (0.033)
S3A_378506623	3A	378506623	C/T	0.17	3.10	0.08	...
S3A_741852990	3A	741852990	A/G	0.08	3.21	0.10	...
S4B_554842477	4B	554842477	A/G	0.34	3.42	0.13	...
S5B_473774480	5B	473774480	A/C	0.28	3.20	0.08	<i>Vrn-A1</i> (0.036)
S5B_501480761	5B	501480761	A/C	0.28	3.28	0.10	...
S5B_502451973	5B	502451973	A/G	0.39	3.33	0.10	...
S5B_503326206	5B	503326206	C/T	0.26	3.07	0.10	...
S5B_504309131	5B	504309131	A/C	0.25	3.72	0.12	...
S5B_508031185	5B	508031185	C/T	0.23	3.08	0.10	...
S5B_509191110	5B	509191110	G/C	0.27	3.00	0.09	...
S5B_513590441	5B	513590441	T/C	0.22	3.89	0.11	...
S5B_528990456	5B	528990456	G/C	0.36	3.85	0.12	...
S7A_483878120	7A	483878120	C/T	0.38	3.01	0.10	...
S7B_749474154	7B	749474154	G/A	0.38	4.05	0.14	...
Severity							
S1A_582293281	1A	582293281	G _r /A _s	0.07	3.22	0.10	<i>Vrn-A1</i> (0.133), <i>Lr34</i> (0.08)
S2A_16824871	2A	16824871	T/C	0.22	3.14	0.10	...
S3A_378506623	3A	378506623	C/T	0.17	3.15	0.08	...
S5A_50162259	5A	50162259	A/G	0.19	3.42	0.11	...
S5B_504309131	5B	504309131	A/C	0.25	3.15	0.10	...
S5B_513590441	5B	513590441	T/C	0.22	3.27	0.08	...
S5B_517797766	5B	517797766	G/C	0.36	3.14	0.08	...
S5B_528990456	5B	528990456	A/G	0.44	3.79	0.12	...
S5B_679369233	5B	679369233	C/T	0.06	3.04	0.08	...
S6B_9296088	6B	9296088	T/C	0.24	3.08	0.12	<i>Vrn-A1</i> (0.043)
S6B_673978653	6B	673978653	C/T	0.07	3.27	0.08	<i>Lr-34</i> (0.049), <i>Vrn-A1</i> (0.066)
S7B_749474154	7B	749474154	G _r /A _s	0.38	3.28	0.11	<i>Vrn-A1</i> (0.057)

^a Abbreviations: Chr = chromosome, SNP = single-nucleotide polymorphism, MAF = minimum allele frequency, and LD = linkage disequilibrium. Markers in bold indicate double association with AUDPC as well as severity.

as *A. squarrosa* (D-genome wheat progenitor). Out of 69, group V and group VI have 25 genotypes with *A. squarrosa* in their pedigree. A subset of synthetic wheat genotypes in group VI was also screened for SB resistance by Iftikhar et al. (2009). Group VI genotypes were basically the synthetic bread wheat derivatives also used in the present study. Due to their similar genetic makeup, these lines were found in the same group differing from others. From this, it can be attributed that some elite cultivars have been more commonly utilized in the pedigrees of the whole collection of studied germplasm.

Factors such as allele frequency, population structure, and recombination rate can affect LD (Flint-Garcia et al. 2003). In the present study, LD decay was observed at a physical distance of approximately 0 to 100 Mb, showing extensive application of markers to scan the high base-pair distance from 0 to 800 Mb pairs. Almost equal proportions of markers in significant LD were observed in the A and B genomes; however, frequency of markers was low in the D genome. One significant advantage of the GBS approach was the availability of a high number of markers in complete LD as compared with the earlier studies; therefore, in the present study, the physical distance has been discussed in megabase pairs rather than the genetic distance in centimorgans used in previous studies (Chen et al. 2016; Gao et al. 2016; Li et al. 2016). Though average extent of LD was lower than found by Rasheed et al. (2014), the percentage of numbers of pairs in complete LD was higher in the present work, indicating increased de novo SNP discovered in the population through GBS.

MTA for SB resistance. All 24 MTA explained phenotypic variation ranging from 8 to 14%, with an average of 10% for SB severity. The higher number of MTA identified in the B-genome corroborates previous studies, where many putative SB resistance genes (*Sb2* and *Sb3*) have been reported (Kumar et al. 2015; Lu et al. 2016). In our study, there are only two markers found associated on the D-genome, although the D-genome is known for having an SB resistance gene (*Sb1*) on 7D (Lillemo et al. 2013). Previously, an SB-associated region on 1A explaining 2.1% variability had been reported by Adhikari et al. (2012) while the present study not only reports the region on 1A elucidating higher (10%) variation in DS but also provides deep insight on this marker (S1A_582293281) by discovering its high LD with a vernalization gene (*Vrn-A1*), with 13% R^2 , and also with gene *Lr34*, with 8% R^2 . It provides a clue for the correlation of SB with these genes but further studies still may be designed to validate this association.

The GBS marker S2A_16824871 on chromosome 2A was associated with severity. Kumar et al. (2009) also identified a QTL, *QSh.bhu-2A*, on chromosome 2AL with an R^2 value of 14.80%, which is likely to be the same QTL. The marker S5A_50162259 harboring a single MTA associated with SB severity explained 10% of the phenotypic variation. Gurung et al. (2014) reported two markers for SB on chromosome 5A that accounted for 22 and 14% of the phenotypic variations, respectively, using the 9K platform. Zhu et al. (2014) reported a QTL (*QSh-cim-5A*) on chromosome 5A derived from synthetic bread wheat. Eleven MTA were identified on chromosome 5B. Previously Kumar et al. (2009) reported a chromosome arm 5BL QTL (*QSh.bhu-5B*) with 38.62% R^2 . Later, Gurung et al. (2014) also confirmed three SNP on 5B associated with SB explaining 17, 58, and 50% of the variation. Due to high values of R^2 , in the recent studies performed by Kumar et al. (2015), an SB resistance gene (*Sb2*) has been claimed on the long arm of 5B.

Two MTA (S6B_9296088 and S6B_673978653) were identified on chromosome 6B, explaining 12 and 8% of the phenotypic variation for severity. Previously, QTL on chromosome 6B using SNP markers (Gurung et al. 2014) have also been reported. A marker on 7B (7SB_749474154) showed association for both AUDPC and severity, with 14 and 11% R^2 , respectively. Adhikari et al. (2012) also declared association of a DArT allele with SB on chromosome 7B. With the 9K genotyping platform, Gurung et al.

(2014) also confirmed the SNP at 7B with 26% R^2 . Five highlighted markers were common in explaining 11% average variability in both AUDPC and severity; therefore, these markers can be used more confidently.

As compared with all previous studies about association mapping of SB, the present study reported GBS markers with highly significant explained variability. These MTA are under significant LD and the highest LD was noted at chromosome 5B that was within 200 Mb distance. Because the GBS platform provides base pair distances, the physical position of the markers (megabase pairs rather than centimorgans) has been highlighted in the study; thus, the markers with significant LD were used for GWAS. In addition to the already reported loci, the present study reported some new MTA on chromosomes 4B and 7A. The marker (S4B_554842477) on 4B and the marker (S7A_483878120) on 7A were not associated with DS but explained 13 and 10% variability in AUDPC, respectively. Because AUDPC is the combined effect of severity recorded at three different stages as compared with DS, which is the severity recorded at a third (GS77) stage, it is not compulsory for these markers to get linked with AUDPC as well as DS simultaneously. These markers can further be validated by utilizing their sequence, having SNP in the middle through in vitro assortment of resistant and susceptible accessions of any other germplasm with already known disease response in the field. Moreover, the sequence of highly associated markers can be used for the development of user-friendly Kompetitive allele-specific polymerase chain reaction (KASP) markers for efficient in vitro screening for SB resistance.

ACKNOWLEDGMENTS

We thank P. St Amand, A. Bernardo, and G. Bai (United States Department of Agriculture–Agricultural Research Service, Hard Winter Wheat Genetics Research Unit, Manhattan, KS 66506, USA) for supervising the performance of the GBS; and S. Asad (CDRI, NARC, Islamabad, Pakistan) for the provision of isolate spray.

LITERATURE CITED

- Adhikari, T. B., Gurung, S., Hansen, J. M., Jackson, E. W., and Bonman, J. M. 2012. Association mapping of quantitative trait loci in spring wheat landraces conferring resistance to bacterial leaf streak and spot blotch. *Plant Genome* 5:1-16.
- Asad, S., Iftikhar, S., Munir, A., and Ahmad, I. 2009. Characterization of *Bipolaris sorokiniana* isolated from different agro-ecological zones of wheat production in Pakistan. *Pak. J. Bot.* 41:301-308.
- Asad, S., Iftikhar, S., Munir, A., Iftikhar, A., and Ayub, N. 2007. Pathogenic diversity in *Bipolaris sorokiniana* isolates collected from different wheat growing areas of the Punjab and NWFP of Pakistan. *Pak. J. Bot.* 39: 2225-2231.
- Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., and Hu, T. T. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465:627-631.
- Bajgain, P., Rouse, M. N., Tsilo, T. J., Macharia, G. K., Bhavani, S., Jin, Y., and Anderson, J. A. 2016. Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. *PLoS One* 11: e0155760.
- Breseghele, F., and Sorrells, M. E. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172: 1165-1177.
- Chen, G., Zhang, H., Deng, Z., Wu, R., Li, D., Wang, M., and Tian, J. 2016. Genome-wide association study for kernel weight-related traits using SNPs in a Chinese winter wheat population. *Euphytica* 212:173-185.
- Earl, D. A. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359-361.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379.
- Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2611-2620.
- Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, I. 2003. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54:357-374.

- Gao, L., Turner, M. K., Chao, S., Kolmer, J., and Anderson, J. A. 2016. Genome wide association study of seedling and adult plant leaf rust resistance in elite spring wheat breeding lines. *PLoS One* 11:1-25.
- Gurung, S., Mamidi, S., Bonman, J. M., Xiong, M., Brown-Guedira, G., and Adhikari, T. B. 2014. Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS One* 9:e108179.
- Iftikhar, S., Asad, S., Mujeeb-Kazi, A., and Ahmad, I. 2009. In vitro screening of synthetic hexaploid wheat lines against *Bipolaris sorokiniana* in Pakistan. *Pak. J. Bot.* 41:1989-2001.
- Jeger, M. 2004. Analysis of disease progress as a basis for evaluating disease management practices. *Annu. Rev. Phytopathol.* 42:61-82.
- Kumar, S., Röder, M. S., Tripathi, S. B., Kumar, S., Chand, R., Joshi, A. K., and Kumar, U. 2015. Mendelization and fine mapping of a bread wheat spot blotch disease resistance QTL. *Mol. Breed.* 35:218.
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. 2009. Mapping of resistance to spot blotch disease caused by *Bipolaris sorokiniana* in spring wheat. *Theor. Appl. Genet.* 118:783-792.
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. 2010. Quantitative trait loci for resistance to spot blotch caused by *Bipolaris sorokiniana* in wheat (*T. aestivum* L.) lines 'Ning 8201' and 'Chirya 3'. *Mol. Breed.* 26:477-491.
- Lan, C., Singh, R., Huerta-Espino, J., Calvo-Salazar, V., and Herrera-Foessel, S. 2014. Genetic analysis of resistance to leaf rust and stripe rust in wheat cultivar Francolin# 1. *Plant Dis.* 98:1227-1234.
- Li, G., Xu, X., Bai, G., Carver, B. F., Hunger, R., Bonman, J. M., and Dong, H. 2016. Genome-wide association mapping reveals novel QTL for seedling leaf rust resistance in a worldwide collection of winter wheat. *Plant Genome* 9.
- Lillemo, M., Joshi, A. K., Prasad, R., Chand, R., and Singh, R. P. 2013. QTL for spot blotch resistance in bread wheat line Saar co-locate to the biotrophic disease resistance loci Lr34 and Lr46. *Theor. Appl. Genet.* 126:711-719.
- Lu, P., Liang, Y., Li, D., Wang, Z., Li, W., Wang, G., and Xie, J. 2016. Fine genetic mapping of spot blotch resistance gene *Sb3* in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* 129:577-589.
- Mascher, M., Wu, S., Amand, P. S., Stein, N., and Poland, J. 2013. Application of genotyping-by-sequencing on semiconductor sequencing platforms: A comparison of genetic and reference-based marker ordering in barley. *PLoS One* 8:e76925.
- Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G. Y., and Myles, S. 2015. LinkImpute: Fast and accurate genotype imputation for non-model organisms. *G3: Genes Genomes Genet.* 5:2383-2390.
- Mujeeb-Kazi, A., Gul, A., Ahmad, I., Farooq, M., Rizwan, S., Bux, H., and Delgado, R. 2007. *Aegilops tauschii*, as a spot blotch (*Cochliobolus sativus*) resistance source for bread wheat improvement. *Pak. J. Bot.* 39:1207-1216.
- Neupane, R., Sharma, R., Duveiller, E., Ortiz-Ferrera, G., Ojha, B., Rosyara, U., and Bhatta, M. 2007. Major gene controls of field resistance to spot blotch in wheat genotypes 'Milan/Shanghai#7' and 'Chirya.3'. *Plant Dis.* 91:692-697.
- Ogbonnaya, F. C., Abdalla, O., Mujeeb-Kazi, A., Kazi, A. G., Xu, S. S., Gosman, N., and Tsujimoto, H. 2013. Synthetic hexaploids: Harnessing species of the primary gene pool for wheat improvement. *Plant Breed. Rev.* 37:35-122.
- Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., Dreisigacker, S., Crossa, J., Sánchez-Villeda, H., Sorrells, M., and Jannink, J.-L. 2012a. Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5:103-113.
- Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. 2012b. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7:e32253.
- Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
- Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R., and He, Z. 2017. Crop breeding chips and genotyping platforms: Progress, challenges, and perspectives. *Mol. Plant* 10:1047-1064.
- Rasheed, A., Xia, X., Ogbonnaya, F., Mahmood, T., Zhang, Z., Mujeeb-Kazi, A., and He, Z. 2014. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biol.* 14:128.
- Roelfs, A. P. 1992. Rust Diseases of Wheat: Concepts and Methods of Disease Management. CIMMYT, Mexico, D.F.
- Saintenac, C., Jiang, D., Wang, S., and Akhunov, E. 2013. Sequence-based mapping of the polyploid wheat genome. *G3: Genes Genomes Genet.* 3: 1105-1114.
- Sharma, R. C., Dubin, H. J., Devkota, R. N., and Bhatta, M. R. 1997. Heritability estimates of field resistance to spot blotch in four spring wheat crosses. *Plant Breed.* 116:64-68.
- Sharma, R. C., Pandey-Chhetri, B., and Duveiller, E. 2006. Heritability estimates of spot blotch resistance and its association with other traits in spring wheat crosses. *Euphytica* 147:317-327.
- Singh, P. K., Zhang, Y., He, X., Singh, R. P., Chand, R., Mishra, V. K., and Islam, R. 2015. Development and characterization of the 4th CSISA-spot blotch nursery of bread wheat. *Eur. J. Plant Pathol.* 143:595-605.
- Vasistha, N. K., Balasubramaniam, A., Mishra, V. K., Chand, R., Srinivasa, J., Yadav, P. S., and Joshi, A. K. 2016. Enhancing spot blotch resistance in wheat by marker-aided backcross breeding. *Euphytica* 207:119-133.
- Villareal, R. L., Mujeeb-Kazi, A., Gilchrist, L. I., and Del Toro, E. 1995. Yield loss to spot blotch in spring bread wheat in warm nontraditional wheat production areas. *Plant Dis.* 79:893-897.
- Wang, R., Leng, Y., Ali, S., Wang, M., and Zhong, S. 2017. Genome-wide association mapping of spot blotch resistance to three different pathotypes of *Cochliobolus sativus* in the USDA barley core collection. *Mol. Breed.* 37:44.
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14:415-421.
- Zhu, Z., Bonnett, D., Ellis, M., Singh, P., Heslot, N., Dreisigacker, S., and Mujeeb-Kazi, A. 2014. Mapping resistance to spot blotch in a CIMMYT synthetic-derived bread wheat. *Mol. Breed.* 34:1215-1228.