

Genome-wide association study and Genomic Prediction of spot blotch disease in wheat (*Triticum aestivum* L.) using genotyping by sequencing

Vipin Tomar

Institute of Advanced Research, Gandhi Nagar, Gujarat, India and Borlaug Institute for South Asia, Ludhiana, Punjab, India

Ravi P Singh

Centro Internacional de Mejoramiento de Maiz y Trigo, Texcoco, Mexico

Jesse Poland

Kansas State University, Manhattan, USA

Daljit Singh

Kansas State University

Arun K Joshi

Borlaug Institute for South Asia (BISA) India

Pawan K Singh

Centro Internacional de Mejoramiento de Maiz y Trigo, Texcoco, Mexico

Pradeep K. Bhati

Borlaug Institute for South Asia (BISA) India

Suneel Kumar

National Bureau of Plant Genetic Resources, New Delhi, India

Mokhlesur Rahman

BARI, Bangladesh

Guriqbal S Dhillon

Thapar University Department of Biotechnology

Budhi S Tiwari

Institute of Advanced Research, Gandhi Nagar, Gujarat, India

Uttam Kumar (✉ kumaruttam@hotmail.com)

Borlaug Institute for South Asia <https://orcid.org/0000-0002-6618-3810>

Research article

Keywords: GWAS, genomics prediction, GBS, spot blotch, BLUP, wheat

Posted Date: February 6th, 2020

DOI: <https://doi.org/10.21203/rs.2.22818/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Spot blotch caused by *Bipolaris sorokiniana* is a major constraint in wheat production in tropics and subtropics. There is limited information available on GWAS and study on genomic prediction is completely lacking. To reveal the genetic markers associated with disease resistance, we performed a genome-wide association study (GWAS) for spot blotch disease in 141 spring wheat lines.

Results Based on the testing under natural infection in three years at hot spots location in Pusa, India and Jamalpur, Bangladesh, the genotypes showed significant genetic variation for disease severity. Using Genotyping-by-Sequencing (GBS) based 18637 polymorphic SNP markers and phenotyping from diverse environments, we identified 23 genomic regions across the genome ($p < 0.001$) on 14 chromosomes associated with disease scores. Consistent with the previous reports, a most stable genomic region on chromosome 2B, 5B and 7D were detected across the environments. The new genomic region on chromosome 3D was also identified. We performed functional annotation with wheat genome assembly annotation (IWGSC Ref Seq v1.0) and identified NBS-LRR and 35 other plant defense-related protein families across multiple chromosome regions. Using a five-fold cross-validation scheme, we observed moderate prediction accuracy for 3 of 4 environments indicated that our model was able to successfully capture the quantitative variation underlying the SB variation in our population.

Conclusions The GWAS based on the phenotypic data from PUSA India and BARI Bangladesh resulted in a total of 23 genomic regions on 14 chromosomes. The new genomic region appeared on chromosome 3D associated with Zinc finger protein that play important role in plant disease resistance. The genomic prediction model for spot blotch disease resistance in wheat was tested and obtained moderate prediction accuracy.

Background

Wheat (*Triticum aestivum* L.) is the major staple for more than 35% of the world's population [1]. The pace of wheat improvement must accelerate to meet the projected global food demand by 2050. Green revolution played a key role in India, Pakistan, Nepal and Bangladesh to ensuring food security in this densely populated region of the world [2]. However, wheat production faces multiple threats via rapidly evolving pathogen variants, pests and increased climate variability, which considerably jeopardize crop productivity growth [3, 4, 5, 6]. Breeding wheat for climatic resilience and disease resistance combined with good agronomy can potentially improve wheat productivity to meet the future food demands [7].

Spot blotch caused by *Bipolaris sorokiniana* is a major constraint in wheat production in tropics and subtropics [8, 9]. The pathogen has a worldwide dispersal, but it is predominantly aggressive under conditions of warm, high relative humidity and temperature associated with imbalanced soil fertility. Yield losses are variable but are important in fields with low inputs and under late-sown conditions [2]. *Bipolaris sorokiniana* act as a causal agent for numerous diseases like head blight, seedling blight, foliar blight/ spot blotch, common root rot and black point of wheat, barley, other small cereal grains and

grasses [10]. However, spot blotch of wheat is considered as one of the most important diseases caused by this pathogen in the mega environments characterized by high temperature (coolest month greater than 17°C) and high humidity [11].

When desired level of resistance to several diseases is required, it is often difficult to achieve through conventional breeding approaches. Disease resistance can be inherited both qualitatively and quantitatively, as is the case in many wheat diseases [12, 13, 14, 15]. The genetic basis of spot blotch resistance has earlier been recognized to eight major quantitative trait loci (QTLs), namely QSb.bhu-2A, QSb.bhu-2B, QSb.bhu-2D, QSb.bhu-3B, QSb.bhu-5B, QSb.bhu-6D, QSb.bhu-7B and QSb.bhu-7D [12, 13]. Sharma et al (2007b) [16] reported three microsatellite markers (Xgwm67, Xgwm570 and Xgwm469) linked with spot blotch resistance. Lr34 and Lr46, the two broadly used genes conferring leaf rust resistance have also been reported to contribute to spot blotch resistance [17]. Lr34 gene is reported to be the major locus on chromosome 7D and explains up to 55% phenotypic variation for spot blotch disease resistance and this locus was given the gene designation Sb1 [17]. During past few years, several QTLs and genetic markers for spot blotch resistance have been identified in multiple studies in winter Wheat [18, 19, 20, 21].

Due to changes in pathogen populations, resistance genes can lose their effectiveness over time. Given these challenges, identification and mapping of novel resistance genes would aid breeding for disease resistance in wheat. One approach to identify spot blotch resistance QTLs is through association mapping. This approach leverages historic recombination and generally have better mapping resolution compared to biparental mapping [22]. A promising strategy to identify QTLs for traits of interest is genome wide association study (GWAS) that takes advantage of decreasing sequencing cost and high throughput genotyping assays [23]. A key approach in GWAS is to have enough genome coverage so that functional alleles will be in linkage disequilibrium (LD) with at least one marker [24]. The association studies for disease resistance including spot blotch have been conducted [25, 26, 27, 28, 29, 30].

Limited research was done where same set is exposed to diverse environments in large geographic area for wheat spot blotch. Therefore, the primary objective of this study was to establish marker–trait associations for spot blotch using genotyping by sequencing (GBS) SNP markers in spring wheat in the South Asian wheat growing regions. The significant SNPs can give insights into the biological function and its relationship with resistance more relevant to the South Asian region. This study aims not only to identify novel QTLs but to validate known genomic loci conferring spot blotch resistance. So far, there is no study on genomic predictions for spot blotch disease resistance in wheat, therefore, we made first attempt to test genomic prediction models across environments.

Results

Spot blotch disease was recorded visually on a scale of zero (no symptoms visible) to 100 (completely susceptible) across three years (2016-17, 2017-18 and 2018-19). The populations displayed significant phenotypic variation for spot blotch resistance with nearly continuous distribution of lines in all

environments (**Fig 1**). The mean disease severity ranged from 8.90 to 31.23 in four environments (2016-17 to 2018-19) including Pusa, India and Bangladesh Agricultural Research Institute (BARI), Bangladesh (**Table 1**). Highest mean disease severity was recorded in BARI 2016-17 (Env4) while lowest was at PUSA 2017-18 (Env2). The analysis of variance revealed highest heritability in Env 4 (0.80) while lowest was in Env 3 (0.50). Based on the combined analysis of all environments, we observed moderate heritability (0.55). There was significant Genotype \times Environment interactions ($P < 0.0001$; **Table 2**).

Genetic linkage mapping

The genetic linkage map was prepared using the most significant SNPs found on 14 hexaploid wheat chromosomes. A total of 70 SNPs were used and clustered in 23 linkage groups (**Fig 2**). A linkage group was considered to be different if the gap between them is more than 10cM on the same chromosome. We observed a maximum of three linkage groups on chromosome 2B and 5B. Similarly, two linkage groups were found on each of 2A, 5A, 7A, 7B, and 7D. Maximum number of significant SNPs (18 SNP markers) were observed on chromosome 2B (**Table 3**).

Principal component analysis

Population structure was determined using genotyping information and principal component analysis (PCA) based approach where genotypes clustered in 12 subgroups. The clustering was done using Ward method in JMP v.13 (**Fig 3**). The Group 1 (G-I) consisted of eight lines including the resistant check HD2733. Maximum number of lines included in G-VII (24 lines) while minimum in G-X and G-XII (6 lines in each). Most of the lines in a group shared alleles descended from common parents. The lines without common parents or less than three sibs per family were classified as 'others'. The largest group (G-VII) consists lines with mixed pedigrees including SAUAL, WBLL#1, Kachu #1, BAV92//IRENA/KAUZ, FRANCOLIN#1, MUCUY and PBW343. The parental lines with TRCH/SRTU//KACHU cross in their genetic backgrounds dominated G-VIII. The parental line AKURI was a most common parent in G-IX and similar in case of G-XII where sister lines dominated the group. The intra-chromosomal LD was calculated as the pairwise marker correlations (R^2) between the GBS markers and plotted against the physical distance for significant marker-trait associations (**Fig S1**).

Marker -Trait associations

The GWAS of spot blotch resistance was performed based on the data collected at adult plant stages. After false discovery rate, a q-value (corrected p value) < 0.05 was used as a threshold to identify significantly associated markers. The GWAS results of spot blotch resistance from the trials conducted at Borlaug Institute for South Asia (BISA) in Pusa, India and BARI in Jamalpur, Bangladesh are given in supplementary tables (**Table S1**). A total of 70 most significant SNPs markers associated with spot blotch disease resistance belong to 23 linkage groups were detected on chromosomes 1A, 1B, 1D, 2A, 2B, 3A, 3B, 3D, 5A, 5B, 6A, 7A, 7B and 7D (**Fig 2**). The phenotypic variation explained by the most significant chromosome region ranged from 6.75% (Env2) to 13.65% (Env1) in three years trials. The phenotypic variation explained in individual environments ranged from 7.37-13.65%, 6.75-9.39%, 7.51-12.72% and

7.60-11.68% in Env1, Env2, Env3 and Env4 respectively (**Table S1**). The largest phenotypic variation explained by the SNP marker located on chromosome 5B in Env1 followed by 2B in Env3 explaining up to 13.65% and 12.72% of phenotypic variation respectively. Most of the SNPs marker regions appeared in more than one environment. We detected seven significant chromosome regions on chromosome 2B while six were on chromosome 5B. A total of 15 SNP markers (S1B_646895451, S2A_31851904, S2B_504717, S2B_525073, S2B_594959, S2B_6253562, S2B_8311062, S2B_90662917, S3B_763230831, S3B_763236179, S3B_763267753, S3B_764192662, S4A_710830493, S6B_719904092 and S7D_181974079) were significant in at least two environments (**Table S1**).

Gene functional annotations

The significant SNPs identified from the GWAS analysis were further studied for the known candidate genes relevant to disease resistance using the recently annotated wheat reference sequence (RefSeq V1.0). We identified NBS-LRR and 35 other plant defense-related protein families across multiple chromosome regions. The significant SNPs S2B_13751999 identified in Env1 on chromosome 2B was located between the GeneIDs, TraesCS2B01G030100 and TraesCS2B01G030200. The latter gene play an important role in the resistance of various plant diseases including the downy mildew [31] by producing RPP13 protein while the former gene is involved in synthesis of lectin-receptor kinase which has an important plant immunity function [32]. Similarly, the SNP S2B_699219601 identified in Env4 belongs to the GeneID, TraesCS2B01G505200 also involve in downy mildew disease resistance response and other diseases (**Table 3**). The SNP S2B_8311062 identified in Env1 and Env3 was located close the geneID, TraesCS2B01G018200, which is involved in *NBS-LRR disease resistance protein* synthesis. Similarly, the SNP S2B_28592818 identified in Env4 located close the geneID, TraesCS2B01G058900 also involved in synthesis NBS-LRR disease resistance protein. The annotation of S5B_683352145 revealed that the gene on chromosome 5B identified in Env4 also involved in synthesis of NBS-LRR disease resistance protein family-1. The SNP S2B_6253562 falls within the GeneID, TraesCS2B01G012400 that encodes Avr9/Cf-9 rapidly elicited protein. The Avr9/Cf-9 protein is involved in early signaling events in the Avr9/Cf-9-dependent plant defense response. The most important SNP S5A_595393566 detected in Env1 on chromosome 5A belongs to the gene TraesCS5A01G402800 which mediates spot blotch (*Bipolaris sorokiniana*) resistance in wheat (**Table 3**). The other SNPs S7B_1020705 and S3B_763230831 associated with *Leucine-rich repeat receptor-like protein kinase* and *transmembrane protein* contribute to *Fusarium* resistance in cereals. The SNP (S5B_233586644; geneID TraesCS5B01G128000) explaining highest phenotypic variance (13.65%) was detected in Env1, produces *signal recognition particle subunit SRP68* which play a crucial role in targeting secretory proteins to the rough endoplasmic reticulum membrane. The second highest phenotypic variance explained by the SNP (S2B_504717) located on chromosome 2B involved in DON resistance through cytochrome 450. The SNP on 1A, 1B, 1D, 2A, 2B, 3A, 3B, 4B, 5A, 5B, 6A and 7B found to be involved directly in disease resistance mechanism (**Table 3**).

To further confirm the results, we looked at the common proteins through gene annotation associated with different SNPs, identified independently in different environments. For example, Zinc finger family protein were associated with the SNP on identified on 2B (S2B_15129248) and 3B (S3B_764747435) in

Env1, 3D (S3D_610628298) in Env2 and 3A in Env3 (S3A_67348475). Similarly, the kinase protein family were associated with the SNPs on chromosome 2B (S2B_13761590), 2A (S2A_708482943) and 5B (S5B_683514734) in Env4. The results based on the gene annotation and synthesized proteins are presented in supplementary **Table S2** and the Vann diagram as **Fig 4**.

Evaluation of Prediction Accuracy

For genomic prediction modeling, the 141 lines were randomly divided into training and testing sets of size 4/5 and 1/5, respectively for each of the four environments. In the initial model training step, both genotypes and the observed phenotypes (i.e., phenotypic BLUPs) were used to estimate the genetic marker effects from the training population. The estimated marker effects were subsequently used to predict the phenotypes of the testing set population. This process was repeated 100 times to sample a random set of training and testing set population during each iteration. The average correlation between the observed and the predicted phenotypes was calculated. The prediction accuracy distributions, means, and standard errors were reported by each environment. The cross-validation prediction accuracy of the Env4 was 0.33 while for the Env1 and Env2 had a prediction accuracy of 0.29 and 0.24 respectively (**Fig 5**). In contrast to other environments, the prediction accuracy of the Env3 was negative (-0.16).

Discussion

The field trials were conducted at BISA research farm, Pusa, in India for three consecutive years from 2016-17 to 2018-19 and BARI farm, Jamalpur in Bangladesh during 2016-17 crop season. Both the locations fall under the non-traditional, warmer wheat-growing regions belonging to Mega-environment 5 characterized by hot, humid conditions as per CIMMYT's system for classifying wheat-growing environments in developing countries [11]. The average temperature during the wheat plant reproductive phase at Jamalpur and Pusa is higher than 19⁰C with a high relative humidity [33] (Table S3).

The spot blotch disease incidence was captured as percentage of infected leaf area at three different growth stages to minimize the chances of disease escape due to environmental factors. However, the scoring date showing highest disease pressure (usually the second one) was used in the analysis. Since the susceptible parent displayed highest disease severity at growth stage 77 (GS77) on Zadoks scale [34], to make better judgment about the level of resistance, disease severity was recorded at this stage (usually second scoring) was used to differentiate each line.

The nearly continuous distribution of lines in all the environments show quantitative nature of resistance. The same has been supported by earlier findings where more than two genes [35, 12, 13, 36, 37] and multiple alleles with minor effect [30, 21] to control spot blotch resistance is reported. It was found that the log transformation improved the data normality which was also reflected by the improved consistency in the GWAS results across locations. We observed significant genetic variation for disease susceptibility in the population. The genetic variances and moderate to high heritabilities for spot blotch were comparable with earlier findings in wheat [37, 38, 36]. Despite significant genotype × environments

interactions, we observed moderate to high heritabilities within environments (Table 2). The environmental interactions might ascribe to difference in the pathogen isolates prevalent in NEPZ of India and Bangladesh in case of locations and weather conditions mainly within location. For example, the maximum mean disease severity of the susceptible lines were up to 43% in Env3 while it was 70% in Env1 (Table 1).

The linkage analysis was based on 18637 filtered SNP markers covering all chromosomes. The redundant SNPs with 0 cM distance and with same gene annotation were removed from the linkage mapping as no additional information is expected. After GWAS analysis, 14 chromosomes harboring significant QTL regions forming 23 linkage groups were used for further analysis and graphical representations. The SNP lies more than 10 cm apart based on linkage mapping, were placed in a separate linkage group. (Fig. 2).

We used genotyping information for the PCA where most of the groups were based on the proportion of genome shared by the parental pool except few exceptions. For example, the subgroup (G-VIII) consisted common parent TRCH/SRTU//KACH while the largest group (G-VII) consists lines with mixed pedigrees dominated by SAUAL, WBLL#1, Kachu #1, BAV92//IRENA/KAUZ, FRANCOLIN#1, MUCUY and PBW343.

Several spot blotch resistance QTLs have been reported on different chromosomes [39, 16, 40, 12, 13, 41, 42, 17, 43, 20, 18, 44, 45]. However, only three major QTLs designated as Sb1 on 7D [17], Sb2 on 5B [41], and Sb3 on 3B [44] are well described. We also observed consistent chromosome regions on 2B and 5B, appeared in more than 2 or all the environments (Table 3). The QTL on 5B, named as Sb2 gene earlier have been studies in detail [44]. The Sb2 gene is known to interact with Tsn1 gene, conferring susceptible reaction to tan spot and Septoria nodorum blotch [46]. The gene ToxA virulent to Tsn1 exists in both *Pyrenophora tritici-repentis* and *Parastagonospora nodorum* confer susceptible reaction to tan spot and Septoria nodorum blotch respectively [21]. Friesen et al. 2018 [47] demonstrated major effects of the Tsn1 locus on chromosome 5B. However, the importance of Tsn1 in spot blotch disease resistance under field condition is not known. The QTL on 7D was the first one studied in detail and reported to be associated with Lr46 [17], Lr34 and leaf tip necrosis [36]. Based on the fine mapping studies, it was named as Sb2 gene [41]. It is interesting to note that Ayana et al. 2018 [30] identified six potential QTLs (Qsb.sdsu-2D.1, Qsb.sdsu-3A.1, Qsb.sdsu- 4A.1, Qsb.sdsu-4B.1, Qsb.sdsu-5A.1, Qsb.sdsu-7B.1) in hard winter wheat using the isolate, SD40 in greenhouse conditions. The chromosome regions on 4A (Env2 and Env4) and 4B (Env4) and 7B (Env1 and Env3) were consistent with the results of [30]. Similarly, four chromosome regions on 1B, 3B, 4B and 5B are validating the finding of [21] which were based on testing in the field condition.

Regardless of % phenotypic variance explained by an allele, almost all wheat chromosomes except 3D and 5D reported to have contributed for spot blotch disease resistance depending on, spot blotch isolate, the breeding material or parents in case of bi-parental population [16, 12, 13, 42, 17, 20, 18, 19, 30, 21]. The minor QTL were reported on, 1BS, 1D, 2D and 3A, 4DS and 6D contributed by 'CIANO T79', 'WUYA' and 'BARTAI' [21]. The broad range of environmental conditions at our field sites allowed us to capture

considerable genetic variation underlying spot blotch resistance. We identified 23 QTL regions on 14 chromosomes validating previous results. The new genomic region detected on chromosome 3D associated with the SNP marker S3D_610628298 explained up to 6.94% of the phenotypic variance but detected in Env2 only. Similarly, the SNP (S2B_90662917) on chromosome 2B was most significant, explained only up to 10% of phenotypic variance, while the SNP on 5B explained largest phenotypic variance in Env1 (Table S1). Out of 23, 9 chromosome regions on seven chromosomes (1A, 1B, 2A, 5A, 5B, 7B, 7D) were already mapped in independent studies earlier [16, 13, 42, 17, 20, 18, 19, 30, 21].

So far, based on the consistency in independent QTL mapping studies using different source of resistance, it seems that there is not much genetic variability in spot blotch pathogen across the continent. However, several studies described clear grouping among spot blotch isolates based on the fungal hyphae color, aggressiveness and DNA fingerprinting [48, 49, 50, 51]. Four chromosomal regions on 1B, 2B, 4A and 6B are consistent between Pusa India and Jamalpur Bangladesh. This may be due to prevalence of most aggressive isolate of spot blotch pathogen (isolate No. ICMP 13584, Auckland, New Zealand) common in South Asia [52].

To study the importance of significant SNPs in disease resistance, we annotated all SNPs using wheat genome assembly annotation (IWGSC Ref Seq v1.0) and traced the protein synthesized by the annotated gene. The literature was mined to look for the putative functions of those proteins. We found that several genes functional annotation strongly associated with disease resistance and observed across the year and environments (Table 3). For example, seven SNPs (S2B_13814702, S2B_533178164, S2B_14809954, S2B_14963432, S2B_15129248, S2B_504717, S2B_78065) on chromosome 2B associated with eight geneIDs, TraesCS2B01G030500, TraesCS2B01G373900, TraesCS2B01G031700, TraesCS2B01G031900, TraesCS2B01G032000, TraesCS2B01G032100, TraesCS2B01G001100 and TraesCS2B01G000400 involved in synthesis of Cytochrome P450 family protein. The role of Cytochrome P450 family protein in plant defense, secondary metabolite biosynthesis in the classical xenobiotic detoxification pathway is well established by Thapa et al. 2018 [53]. It is involved in resistance to DON which is a trichothecene mycotoxin produced by *Fusarium* species and increase yield. The Cytochrome P450 family protein may not involve directly in yield increase but to enhanced *Fusarium* head blight disease resistance.

The SNP (S2B_28592818) detected in Env4 on same chromosome (2B) but at different region synthesizes NBS-LRR disease resistance protein family contribute for disease resistance [54, 55]. Similarly, the SNP S2B_8311062 and S5B_683352145 also associated with the gene synthesize NBS-LRR disease resistance protein family and contribute for fungal disease resistance. The role of NBS-LRR disease resistance protein is disease resistance mechanism is well established [54, 55]. One of the significant SNP located on chromosome 2B, S2B_15129248 is linked to two geneIDs, namely, TraesCS2B01G032100 (synthesize Cytochrome P450 family proteins) and TraesCS2B01G032200 (involved in GRF zinc finger family protein). Both the proteins play an important role in plant disease resistance [53, 56].

It is interesting to note that the most important SNP S5A_595393566 detected in Env1 on chromosome 5A belongs to the gene TraesCS5A01G402800 which mediates spot blotch resistance in wheat. This gene is involved in the synthesis of Myb family transcription factor-like protein, found to mediate host resistance to *Bipolaris sorokiniana* in wheat [57]. The same region has been reported in other independent studies as well [18, 20, 30, 21]. Similarly, the SNP S3A_67065083 associated with geneID TraesCS3A01G103500 involved in synthesis of 1R-MYB Transcription factor which plays an important role in disease resistance against stripe rust fungus and ear head disease in wheat [45].

The key SNPs on chromosome 1A (MAPK module FgSte50-Ste11-Ste7 in *F. graminearum*), 1B (stripe rust & powdery mildew), 1D (Serine/threonine-protein kinase), 2B (RPP13, Avr9/Cf-9 rapidly elicited protein, NBS-LRR protein, F-box family protein, pentatricopeptide repeat-containing protein, Peptidylprolyl isomerase, Uroporphyrinogen decarboxylase and resistance to DON), 3A and 3B (1R-MYB TF, wheat NAC protein and interaction with an orphan protein), 4B (Uroporphyrinogen decarboxylase), 5A (Myb family transcription factor-like, Serine/arginine repetitive matrix, NBS-LRR & transmembrane protein), 5B (B3 domain-containing, Mannitol transporter & NBS-LRR family-1 protein) and 7D (implicated in the defense through cell wall modification, degradation, carbohydrate metabolic processes) annotated and found to synthesize different proteins involved in fungal defense mechanism (Table 3) [58, 59, 60, 55, 61, 54, 62, 63, 64, 65, 32, 66, 67, 68, 69, 57, 70, 71]. The consistency in identification of key SNPs involved in resistance mechanism through protein annotation was confirmed where same protein family was identified independently in different environments (Fig. 4).

Maximum number of known proteins involved in fungal defense were based on 14 SNPs on chromosome 2B showing the importance of this chromosome in disease resistance. The earlier independent findings also describe the importance of chromosome 2B in spot blotch disease resistance [12, 19]. Interestingly QTL found in the present study on 1B in Env1, the proteins involved are Pentatricopeptide repeat-containing protein (TraesCS1B01G424000) mRNAs renders more susceptible to pathogenic bacteria and fungi in *Arabidopsis thaliana* [109] and Homeobox protein (TraesCS1B01G424100) associated with reaction to stripe rust and powdery mildew in common wheat [72]. The SNPs (S1B_646895451 and S1B_647195634) detected in two environments located on chromosome 1B are 18.97 cM apart on genetic linkage map while those belong to the same geneID TraesCS1B01G424000. The gene annotation results indicate role in plant disease resistance [69, 65, 72]. This information obtained from gene annotation could potentially be used in fine mapping and map-based cloning to further characterize the mechanisms of spot blotch disease resistance. The markers with lowest P-values may be converted in to diagnostic markers to validate the SNPs and used in identification of lines with desired alleles in early generations.

Genomic Prediction

We used the rrBLUP GS model, which includes all marker information to predict a line's genomic estimated breeding values (GEBV) [73, 74]. rrBLUP requires much lower computational time compared to the other GS models and it is shown to be robust in different GS scenarios [75, 76, 77].

Improving disease resistance in crops via single or a few major genes may be a temporary solution because the effectiveness is limited to only selected races of the pathogen and therefore, have no broad-environment application [78, 79]. The costs connected with the introgression of major genes or QTLs into elite backgrounds is challenging and may unintentionally affect the breeding operations and fast-track the evolution of the pathogen. Here, genomic prediction is well-placed to improve the effectiveness of quantitative disease resistance efforts in wheat breeding [80, 81]. The SB in wheat is shown to be controlled by many small to large effect genes [12, 13, 82, 83]. Therefore, instead of focusing on a few large-effect QTL, the prediction of spot blotch disease infection based on both small- and large-effect QTLs promise a holistic genetic-based approach for broad-spectrum resistance to evade the development and spread of spot blotch contagion.

Our results provide additional evidence in support of the quantitative nature of the disease resistance in wheat. Moderate to high prediction accuracies for 3 of 4 environments indicated that our model was able to successfully capture the quantitative variation underlying the SB variation in our population (**Fig. 5**). The contrasting prediction accuracies for PUSA19 environments in our study underscore the need for additional research to investigate the stability of genomic predictions across environments [75, 84, 85]. The role of environmental variation in the form of genotype-by-environment and marker-by-environment interactions in genomic selection has been highlighted by others [86, 87]. In view of both the positive and negative findings in this genomic prediction work, this study would provide an important precursor for future wheat breeding research in South Asia which is proven by other researchers also [79, 81].

Conclusions

This study aimed to identify genetic regions underlying spot blotch resistance in the elite spring wheat genotypes. The variable conditions at four field environments in India and Bangladesh allowed us to capture the considerable phenotypic variation for spot blotch disease in our trials. The GWAS based on the phenotypic data at each site resulted in a total of 23 genomic regions on 14 chromosomes. We were able to validate earlier findings and identified new genomic regions on chromosome 3D contributing up to 6.94% of phenotypic variation. The literature mining of the functional gene annotations identified 36 SNPs encoding single protein or protein family, directly or indirectly involved in disease resistance. The SNPs on chromosome 5A associated with the known gene encodes 'Myb family transcription factor-like protein' found to have direct involvement in spot blotch resistance. Using a five-fold cross-validation scheme, we observed moderate prediction accuracies for 3 of 4 environments indicated that our model was able to successfully capture the quantitative variation underlying the SB variation in our population. The results are of importance for the breeders developing spot blotch resistant varieties targeting South Asian region. Given the aggressive pathogen spread and food security concerns, the breeding programs in South Asia could benefit by deploying a genomic selection based breeding scheme for broad-spectrum spot blotch disease resistance in wheat.

Methods

Plant material and field layout

The population was a genetically diverse collection comprising 141 advanced breeding lines of spring wheat. It represents 25 years of research at CIMMYT and was carefully assembled to avoid the confounding effects of phenology. The lines were evaluated in two replicates at two field locations: BISA, Pusa, India (25°57'22.8"N, 85°40'13.1"E) in the north India, and BARI, Jamalpur, Bangladesh (24°22'07.7"N, 88°39'42.0"E) during 2016-17, 2017-18 and 2018-19 seasons in a Randomized Block Design. For convenience, the different location-season combinations were termed as Env1 (Pusa 2016-17), Env2 (Pusa 2017-18), Env3 (Pusa 2018-19) and Env4 (BARI 2016-17). Both these locations are known as hot spot for spot blotch disease [3]. The plots of 3.8-meter length were sown in six rows with 0.22 meter spacing at each environment. The trials were timely sown with full irrigation applied through gravity flood-irrigation. The spreader rows of susceptible variety Sonalika were also planted in alleys for disease build up. In addition, four auxiliary gravity flood-irrigations were also given at regular intervals. All agronomic practices like fertigation and weeding were performed as recommended for each location.

Screening for spot blotch disease resistance

The material was evaluated under natural infection conditions in the field. To limit the number of escapes, spot blotch response was evaluated during the mid to advanced-phases of disease development *i.e.*, between heading (GS50 on Zadoks scale) and grain filling stage (GS80) [34]. The disease severity (SEV) were recorded visually on 0 to 100 scale where 0 is complete resistance and 100 is complete susceptibility.

Genotyping

Seeds of all lines were obtained from the CIMMYT genetic resources program and genomic DNA was extracted from five bulked leaves using a CTAB procedure [88,] modified as described in Dreisigacker et al. 2013 [89] in CIMMYT, Mexico. The DNA samples were sent to Kansas State University, USA for GBS. The GBS was performed following the protocol of Poland et al 2012 [90]. All lines were sequenced with Illumina Hi Seq2000 or HiSeq2500. GBS-SNP markers were called with TASSEL v5.2 pipeline [91] and aligned to the reference Chinese Spring Wheat Assembly v1.0. The following SNP filtering criteria was applied on raw SNP calls: less than 30% missing markers, minimum 5% minor allele frequency (MAF) and less than 20% heterozygosity. The filtering step yielded 18637 markers and the remaining missing values were imputed using Beagle v4.1 [92].

Statistical analysis

The experimental design in each environment was an randomized complete block design with two replications per location. META-R v6.03 developed by CIMMYT [93], Mexico was used to perform multi-environment mixed model analysis. The environments were used as random effects and genotypes as fixed effects. The resulting analysis produced the adjusted trait phenotypic values in the form of best linear unbiased predictions (BLUPs) within and across environments. In addition, the components of

phenotypic variance were also extracted to calculate broad-sense line-mean heritability. The genetic linkage map was prepared using IciMapping v.4.0 [94] while principal component and dendograms were performed using JMP v13 (SAS Inc., Cary, NC, USA). In addition to BLUP, we calculated Log, square root and Arrhenius values from severity percentage using JMP v13. The raw phenotypic distributions of disease scores at each environment were plotted to check normality assumptions. The data was log-transformed for cases where the distributions deviated significantly from normality.

PCA and linkage disequilibrium

The PCA was performed using 18637 SNPs and 141 genotypes in Tassel v5.2 [91] for the genetic relationship among genotypes. The first two principal components were drawn to show the clustering among genotypes. The population structure was inferred using the JMP v13. The Kinship was calculated using IBS method in Tassel 5.2. In order to determine the number of SNP marker for GWAS, the LD was estimated in TASSELv5.2 [91] using 18637 markers. The long-distance LD approach was used where critical r^2 of 0.2 was drawn in R software v3.5.2. [95] The marker r^2 for some chromosomes showed extensive noise, therefore instead of calculating chromosome-wide LD-decay thresholds, we generated LD heat maps of the significant markers for each chromosome separately. The intra-chromosomal LD was calculated as the pairwise marker correlations (R^2) between the GBS markers and plotted against the physical distance for significant marker-trait associations.

Association analysis

TASSEL v5.2 [91] was used to calculate population Kinship Matrix based on scaled IBS (identity by state) method. Mixed linear model (MLM) was used to test the marker–trait association between the SNP markers and spot blotch disease severity (SEV). MLM has proven useful in controlling for population structure and relatedness within genome wide association studies. Subsequent GWAS and genomic prediction analyses were performed on 16037 SNPs scored on 141 lines from the seasons 2016-17, 2017-18 and 2018-19. Genome-wide significance threshold was established to reduce the false discovery rate due to multiple test comparisons.

Gene functional annotations

GWAS results were further analyzed to test if the marker-trait associations fall within the known genic regions and by functional annotation from the reference genome assembly (IWGSC Ref Seq v1.0). Functional annotation of the genes harboring significant SNPs were retrieved and examined for their association with spot blotch resistance from the genome annotations provided by IWGSC. Subsequently, genes annotated proteins functions were literature mined.

Genomic Prediction

The Log transformed BLUPs of spot blotch scores calculated across years were used for genomic prediction modeling. A five-fold cross-validation scheme [96] was implemented. The advanced breeding

lines were randomly divided into five subsets (i.e., folds), and four of them were used as the training set. each with approximately the same number of individuals. The random cross-validation runs were repeated for 100 iterations. At each step, the predictive accuracy of the markers was assessed by Pearson's correlation between the predicted values and the BLUP phenotypes. Overall average of the fifth fold was reported as accuracy of the prediction. All calculations were performed in R software [95] and by using the packages *lme4* and *rrBLUP* [97, 74).

Declarations

Acknowledgements

This work was supported by the US Agency for International Development (USAID) Feed the Future Innovation Lab for Applied Wheat Genomics (Cooperative Agreement No. AID-OAA-A-13-00051). The technical support of Manish Kumar and Raj Kumar Jat is dully acknowledged.

Author contribution

VT: Drafted the manuscript, recorded data in the field and analyzed the data; **RPS:** Provided the breeding material and experiment design; **JP:** Genomic predictions and GBS data of breeding material; **DJ:** Performed genomic predictions; **AKJ:** Spot evaluation and trial management; **PKS:** scientific inputs for manuscript preparation; **PKB:** Experimental design and data recording; **SK:** Recorded data in the field; **MR:** Spot evaluation and trial management at BARI; **GSD:** Linkage mapping analysis; **BST:** revised the manuscript; **UK:** Design the experiment, revised manuscript and overall supervised the research work. All authors read and approved the final manuscript.

Corresponding authors

Correspondence to [Uttam Kumar \(u.kumar@cgiar.org\)](mailto:u.kumar@cgiar.org)

Ethics declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declared that they had no competing interests

Declaration

The data can be made available on request. It is confirmed that the data may be uploaded in to public domain once the manuscript is accepted for publication.

References

1. FAO (2013) FAO statistical year book 2013: World food and agriculture. Rome, Italy.2013
2. Joshi AK, Mishra B, Chatrath R, Ferrara GO, Singh RP. Wheat improvement in India: present status, emerging challenges and future prospects. *Euphytica*. 2007 Oct 1;157(3):431-46.
3. Sharma RC, Duveiller E, Ortiz-Ferrara G. Progress and challenge towards reducing wheat spot blotch threat in the Eastern Gangetic Plains of South Asia: is climate change already taking its toll?. *Field Crops Research*. 2007a Aug 30;103(2):109-18.
4. Gurung S, Bonman JM, Ali S, Patel J, Myrfield M, Mergoum M, Singh PK, Adhikari TB. New and diverse sources of multiple disease resistance in wheat. *Crop science*. 2009 Sep 1;49(5):1655-66.
5. Gurung S, Hansen JM, Bonman JM, Gironella AI, Adhikari TB. Multiple disease resistance to four leaf spot diseases in winter wheat accessions from the USDA National Small Grains Collection. *Crop science*. 2012 Jul 1;52(4):1640-50.
6. Friesen TL, Xu SS, Harris MO. Stem rust, tan spot, *Stagonospora nodorum* blotch, and Hessian fly resistance in Langdon durum–*Aegilops tauschii* synthetic hexaploid wheat lines. *Crop science*. 2008 May 1;48(3):1062-70.
7. Mondal S, Rutkoski JE, Velu G, Singh PK, Crespo-Herrera LA, Guzman C, Bhavani S, Lan C, He X, Singh RP. Harnessing diversity in wheat to enhance grain yield, climate resilience, disease and insect pest resistance and nutrition through conventional and modern breeding approaches. *Frontiers in plant science*. 2016 Jul 6;7:991.
8. Dubin HJ, Van Ginkel M. The status of wheat diseases and disease research in the warmer areas. In: Saunder DA, Hettel GP, editors. *Wheat for non-traditional warm areas in heat stressed environments: irrigated dry areas*. Mexico: UNDP/CIMMYT; 1991;p.125–145.
9. Saari EE. Leaf blight disease and associated soil-borne fungal pathogens of wheat in South and South East Asia. *Helminthosporium* Blights of wheat: spot blotch and tan spot. 1998:37-51.
10. Zillinsky FJ. Common diseases of small grain cereals. A guide to identification. *Common diseases of small grain cereals. A guide to identification..* 1983.
11. Van Ginkel M, Rajaram S Breeding for durable resistance to diseases in wheat : an additional perspective. In: Jacobes T, Parlevliet JE (eds) *Durability of disease resistance*. Kluwer Academic Publishers, Dordrecht, Netherlands, 1993. pp 259-272

12. Kumar U, Joshi AK, Kumar S, Chand R, Röder MS. Mapping of resistance to spot blotch disease caused by *Bipolaris sorokiniana* in spring wheat. *Theoretical and Applied Genetics*. 2009 Feb 1;118(4):783-92.
13. Kumar U, Joshi AK, Kumar S, Chand R, Röder MS. Quantitative trait loci for resistance to spot blotch caused by *Bipolaris sorokiniana* in wheat (*aestivum* L.) lines 'Ning 8201' and 'Chirya 3'. *Molecular Breeding*. 2010 Oct 1;26(3):477-91.
14. Marone D, Russo MA, Laidò G, De Vita P, Papa R, Blanco A, Gadaleta A, Rubiales D, Mastrangelo AM. Genetic basis of qualitative and quantitative resistance to powdery mildew in wheat: from consensus regions to candidate genes. *BMC genomics*. 2013 Dec;14(1):562.
15. Elizabeth F, Kim BS, Iyer-Pascuzzi AS. Mechanisms of quantitative disease resistance in plants. *Seminars in Cell & Developmental Biology* 2016;56:201-208
16. Sharma RC, Duveiller E, Jacquemin JM. Microsatellite markers associated with spot blotch resistance in spring wheat. *Journal of phytopathology*. 2007b May;155(5):316-9.9
17. Lillemo M, Joshi AK, Prasad R, Chand R, Singh RP. QTL for spot blotch resistance in bread wheat line Saar co-locate to the biotrophic disease resistance loci *Lr34* and *Lr46*. *Theoretical and applied genetics*. 2013 Mar 1;126(3):711-9.
18. Zhu X, Qi L, Liu X, Cai S, Xu H, Huang R, Li J, Wei X, Zhang Z. The wheat ethylene response factor transcription factor pathogen-induced ERF1 mediates host responses to both the necrotrophic pathogen *Rhizoctonia cerealis* and freezing stresses. *Plant physiology*. 2014 Mar 1;164(3):1499-514.
19. Singh G, Sheoran S, Chowdhury AK, Tyagi BS, Bhattacharya PM, Singh V, Ojha A, Turan R, Sharma I. Phenotypic and marker aided identification of donors for spot blotch resistance in Wheat. *Journal of Wheat Research*. 2014;6(1).
20. Gurung S, Mamidi S, Bonman JM, Xiong M, Brown-Guedira G, Adhikari TB. Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS One*. 2014 Sep 30;9(9):e108179.
21. Singh P, He X, Sansaloni C, Juliana P, Dreisigacker S, Duveiller E, Kumar U, Joshi A, Singh R. Resistance to Spot Blotch in Two Mapping Populations of Common Wheat Is Controlled by Multiple QTL of Minor Effects. *International journal of molecular sciences*. 2018 Dec;19(12):4054.
22. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. *The plant genome*. 2008 Jul 1;1(1):5-20.
23. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS One* 2011;6:e19379.
24. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell*. 2009 Aug 1;21(8):2194-202.
25. Maccaferri M, Sanguineti MC, Mantovani P, Demontis A, Massi A, Ammar K, Kolmer JA, Czembor JH, Ezrati S, Tuberosa R. Association mapping of leaf rust response in durum wheat. *Molecular Breeding*. 2010 Aug 1;26(2):189-228.

26. Maccaferri M, Zhang J, Bulli P, Abate Z, Chao S, Cantu D, Bossolini E, Chen X, Pumphrey M, Dubcovsky J. A genome-wide association study of resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*) in a worldwide collection of hexaploid spring wheat (*Triticum aestivum*). *G3: Genes, Genomes, Genetics*. 2015 Mar 1;5(3):449-65.
27. Yu LX, Lorenz A, Rutkoski J, Singh RP, Bhavani S, Huerta-Espino J, Sorrells ME. Association mapping and gene–gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theoretical and applied genetics*. 2011 Dec 1;123(8):1257-68.
28. Zhang B, Liu X, Zhao G, Mao X, Li A, Jing R. Molecular characterization and expression analysis of *Triticum aestivum* squamosa-promoter binding protein-box genes involved in ear development. *Journal of integrative plant biology*. 2014 Jun;56(6):571-81.
29. Zegeye H, Rasheed A, Makdis F, Badebo A, Ogonnaya FC. Genome-wide association mapping for seedling and adult plant resistance to stripe rust in synthetic hexaploid wheat. *PLoS One*. 2014 Aug 25;9(8):e105593.
30. Ayana GT, Ali S, Sidhu JS, Gonzalez-Hernandez JL, Turnipseed B, Sehgal SK. Genome-wide association study for spot blotch resistance in hard winter wheat. *Frontiers in Plant Science*. 2018;9:926.
31. Cheng J, Fan H, Li L, Hu B, Liu H, Liu Z. Genome-wide Identification and Expression Analyses of RPP13-like Genes in Barley. *BioChip Journal*. 2018 Jun 1;12(2):102-13.
32. Wang Y, Bouwmeester K, Beseh P, Shan W, Govers F. Phenotypic analyses of Arabidopsis T-DNA insertion lines and expression profiling reveal that multiple L-type lectin receptor kinases are involved in plant immunity. *Molecular Plant-Microbe Interactions*. 2014 Dec;27(12):1390-402.
33. JAMALPUR WEATHER BY MONTH // WEATHER AVERAGES: <https://en.climate-data.org/asia/bangladesh/dhaka-division/jamalpur-57280/>. Accessed on June 2019
34. Zadoks JC, Chang TT, Konzak CF. A decimal code for the growth stages of cereals. *Weed research*. 1974 Dec;14(6):415-21.
35. Kumar S, Prasad LC, Kumar U, Tyagi K, Arun B, Joshi AK. Inheritance and allelic relationship of resistance genes to spot blotch of wheat caused by *Bipolaris sorokiniana*. In *Wheat Production in Stressed Environments 2007* (pp. 113-118). Springer, Dordrecht.
36. Joshi AK, Kumar S, Chand R, Ortiz-Ferrara G. Inheritance of resistance to spot blotch caused by *Bipolaris sorokiniana* in spring wheat. *Plant Breeding*. 2004 Jun;123(3):213-9.
37. Kumar S, Kumar U, Singh AK, Singh R, Prasad LC, Joshi AK. Study of inheritance and allelic relation of resistance to spot blotch (*Bipolaris sorokiniana*) of wheat. *Biotech Today*. 2013;3(2):31-8.
38. Sharma RC, Dubin HJ, Devkota RN, Bhatta MR. Heritability estimates of field resistance to spot blotch in four spring wheat crosses. *Plant Breeding*. 1997 Mar;116(1):64-8.
39. Neupane RB, Sharma RC, Duveiller E, Ortiz-Ferrara G, Ojha BR, Rosyara UR, Bhandari D, Bhatta MR. Major gene controls of field resistance to spot blotch in wheat genotypes 'Milan/Shanghai# 7' and 'Chirya. 3'. *Plant disease*. 2007 Jun;91(6):692-7.

40. Gonzalez-Hernandez JL, Singh PK, Mergoum M, Adhikari TB, Kianian SF, Simsek S, Elias EM. A quantitative trait locus on chromosome 5B controls resistance of *Triticum turgidum* (L.) var. *diccocoides* to *Stagonospora nodorum* Euphytica. 2009 Mar 1;166(2):199.
41. Kumar S, Röder MS, Tripathi SB, Kumar S, Chand R, Joshi AK, Kumar U. Mendelization and fine mapping of a bread wheat spot blotch disease resistance QTL. Molecular breeding. 2015 Nov 1;35(11):218.
42. Adhikari TB, Gurung S, Hansen JM, Jackson EW, Bonman JM. Association mapping of quantitative trait loci in spring wheat landraces conferring resistance to bacterial leaf streak and spot blotch. The Plant Genome. 2012 Mar 1;5(1):1-6.
43. Zhuang Y, Gala A, Yen Y. Identification of functional genic components of major Fusarium head blight resistance quantitative trait loci in wheat cultivar Sumai 3. Molecular plant-microbe interactions. 2013 Apr;26(4):442-50.
44. Lu P, Liang Y, Li D, Wang Z, Li W, Wang G, Wang Y, Zhou S, Wu Q, Xie J, Zhang D. Fine genetic mapping of spot blotch resistance gene Sb3 in wheat (*Triticum aestivum*). Theoretical and applied genetics. 2016 Mar 1;129(3):577-89.
45. Zhang Z, Chen J, Su Y, Liu H, Chen Y, Luo P, Du X, Wang D, Zhang H. TaLHY, a 1R-MYB transcription factor, plays an important role in disease resistance against stripe rust fungus and ear heading in wheat. PLoS One. 2015 May 26;10(5):e0127723.
46. McDonald MC, Ahren D, Simpfendorfer S, Milgate A, Solomon PS. The discovery of the virulence gene ToxA in the wheat and barley pathogen *Bipolaris sorokiniana*. Molecular plant pathology. 2018 Feb;19(2):432-9.
47. Friesen TL, Holmes DJ, Bowden RL, Faris JD. ToxA is present in the US *Bipolaris sorokiniana* population and is a significant virulence factor on wheat harboring Tsn1. Plant disease. 2018 Dec 16;102(12):2446-52.
48. Chand R, Pandey SP, Singh HV, Kumar S, Joshi AK. Variability and its probable cause in natural populations of spot blotch pathogen *Bipolaris sorokiniana* of wheat (*T. aestivum* L.) in India. J. Plant Dis. Protect. 2003;110:27-35.
49. Jaiswal SK, Prasad LC, Sharma S, Kumar S, Prasad R, Pandey SP, Chand R, Joshi AK. Identification of molecular marker and aggressiveness for different groups of *Bipolaris sorokiniana* isolates causing spot blotch disease in wheat (*Triticum aestivum*). Current Microbiology. 2007 Aug 1;55(2):135-41.
50. Mahto BN, Gurung S, Nepal A, Adhikari TB. Morphological, pathological and genetic variations among isolates of *Cochliobolus sativus* from Nepal. European journal of plant pathology. 2012 Jun 1;133(2):405-17.
51. Sultana S, Adhikary SK, Islam MM, Rahman SM. Evaluation of pathogenic variability based on leaf blotch disease development components of *Bipolaris sorokiniana* in *Triticum aestivum* and agroclimatic origin. The plant pathology journal. 2018 Apr;34(2):93.

52. Chaurasia S, Chand R, Joshi AK. Relative dominance of *Alternaria triticina* et Prab. and *Bipolaris sorokiniana* (Sacc.) Shoemaker in different growth stages of wheat (*T. aestivum* L.). *Journal of Plant Diseases and Protection*. 2000 Mar 1;176-81.
53. Thapa G, Gunupuru LR, Hehir JG, Kahla A, Mullins E, Doohan FM. A pathogen-responsive leucine rich receptor like kinase contributes to Fusarium resistance in cereals. *Frontiers in plant science*. 2018 Jun 26;9:867.
54. Liu W, Frick M, Huel R, Nykiforuk CL, Wang X, Gaudet DA, Eudes F, Conner RL, Kuzyk A, Chen Q, Kang Z. The stripe rust resistance gene Yr10 encodes an evolutionary-conserved and unique CC–NBS–LRR sequence in wheat. *Molecular plant*. 2014 Dec 1;7(12):1740-55.
55. He H, Zhu S, Zhao R, Jiang Z, Ji Y, Ji J, Qiu D, Li H, Bie T. Pm21, encoding a typical CC-NBS-LRR protein, confers broad-spectrum resistance to wheat powdery mildew disease. *Molecular plant*. 2018 Jun 4;11(6):879-82.
56. Gupta SK, Rai AK, Kanwar SS, Sharma TR. Comparative analysis of zinc finger proteins involved in plant disease resistance. *PLoS One*. 2012 Aug 15;7(8):e42578.
57. Zhang Z, Liu X, Wang X, Zhou M, Zhou X, Ye X, Wei X. An R2R3 MYB transcription factor in wheat, Ta PIMP 1, mediates host resistance to *Bipolaris sorokiniana* and drought stresses through regulation of defense-and stress-related genes. *New Phytologist*. 2012 Dec;196(4):1155-70.
58. Gou M, Su N, Zheng J, Huai J, Wu G, Zhao J, He J, Tang D, Yang S, Wang G. An F-box gene, CPR30, functions as a negative regulator of the defense response in Arabidopsis. *The Plant Journal*. 2009 Dec;60(5):757-70.
59. Yang Y, Zhao J, Liu P, Xing H, Li C, Wei G, Kang Z. Glycerol-3-phosphate metabolism in wheat contributes to systemic acquired resistance against *Puccinia striiformis* sp. *tritici*. *PloS one*. 2013 Nov 29;8(11):e81756.
60. Yang Y, Yu Y, Bi C, Kang Z. Quantitative proteomics reveals the defense response of wheat against *Puccinia striiformis* sp. *tritici*. *Scientific reports*. 2016 Sep 28;6:34261.
61. Romeis T, Tang S, Hammond-Kosack K, Piedras P, Blatt M, Jones JD. Early signalling events in the Avr9/Cf-9-dependent plant defence response. *Molecular Plant Pathology: Original Articles*. 2000 Jan;1(1):3-8.
62. Fan X, Naz M, Fan X, Xuan W, Miller AJ, Xu G. Plant nitrate transporters: from gene function to application. *Journal of experimental botany*. 2017 May 1;68(10):2463-75.
63. Liu J, Zhang T, Jia J, Sun J. The wheat mediator subunit TaMED25 interacts with the transcription factor TaEIL1 to negatively regulate disease resistance against powdery mildew. *Plant physiology*. 2016 Mar 1;170(3):1799-816.
64. Fonseca JP, Dong X. Functional characterization of a Nudix hydrolase AtNUDX8 upon pathogen attack indicates a positive role in plant immune responses. *PloS one*. 2014 Dec 1;9(12):e114119.
65. Park YJ, Lee HJ, Kwak KJ, Lee K, Hong SW, Kang H. MicroRNA400-guided cleavage of pentatricopeptide repeat protein mRNAs renders *Arabidopsis thaliana* more susceptible to pathogenic bacteria and fungi. *Plant and Cell Physiology*. 2014 Jul 8;55(9):1660-8.

66. Cao A, Xing L, Wang X, Yang X, Wang W, Sun Y, Qian C, Ni J, Chen Y, Liu D, Wang X. Serine/threonine kinase gene *Stpk-V*, a key member of powdery mildew resistance gene *Pm21*, confers powdery mildew resistance in wheat. *Proceedings of the National Academy of Sciences*. 2011 May 10;108(19):7727-32.
67. Noutoshi Y, Okazaki M, Kida T, Nishina Y, Morishita Y, Ogawa T, Suzuki H, Shibata D, Jikumaru Y, Hanada A, Kamiya Y. Novel plant immune-priming compounds identified via high-throughput chemical screening target salicylic acid glucosyltransferases in *Arabidopsis*. *The Plant Cell*. 2012 Sep 1;24(9):3795-804.
68. Oide S, Liu J, Yun SH, Wu D, Michev A, Choi MY, Horwitz BA, Turgeon BG. Histidine kinase two-component response regulator proteins regulate reproductive development, virulence, and stress responses of the fungal cereal pathogens *Cochliobolus heterostrophus* and *Gibberella zeae*. *Eukaryotic cell*. 2010 Dec 1;9(12):1867-80.
69. Gu Q, Chen Y, Liu Y, Zhang C, Ma Z. The transmembrane protein *FgSht1* regulates fungal development and pathogenicity via the MAPK module *Ste50-Ste11-Ste7* in *Fusarium graminearum*. *New Phytologist*. 2015 Apr;206(1):315-28.
70. Zhang J, Wang F, Liang F, Zhang Y, Ma L, Wang H, Liu D. Functional analysis of a pathogenesis-related thaumatin-like protein gene *TaLr35PR5* from wheat induced by leaf rust fungus. *BMC plant biology*. 2018 Dec;18(1):76.
71. Upadhyay RS, Meena M, Prasad V, Zehra A, Gupta VK. Mannitol metabolism during pathogenic fungal–host interactions under stressed conditions. *Frontiers in microbiology*. 2015 Sep 24;6:1019.
72. Liu D, Xia XC, He ZH, Xu SC. A novel homeobox-like gene associated with reaction to stripe rust and powdery mildew in common wheat. *Phytopathology*. 2008 Dec;98(12):1291-6.
73. Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001 Apr 1;157(4):1819-29.
74. Endelman JB. Ridge regression and other kernels for genomic selection with R package *rrBLUP*. *Plant Genome* 2011;4:250–255.
75. Heslot N, Akdemir D, Sorrells ME, Jannink JL. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theoretical and applied genetics*. 2014 Feb 1;127(2):463-80.
76. Resende MF, Muñoz P, Resende MD, Garrick DJ, Fernando RL, Davis JM, Jokela EJ, Martin TA, Peter GF, Kirst M. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda*). *Genetics*. 2012 Apr 1;190(4):1503-10.
77. Riedelsheimer C, Technow F, Melchinger AE. Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. *BMC genomics*. 2012 Dec;13(1):452.
78. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proceedings of the National Academy of Sciences*. 2011 Apr 26;108(17):6893-8.

79. Singh RP, Hodson DP, Jin Y, Lagudah ES, Ayliffe MA, Bhavani S, Rouse MN, Pretorius ZA, Szabo LJ, Huerta-Espino J, Basnet BR. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology*. 2015 Jun 29;105(7):872-84.
80. Crossa J, de Los Campos G, Pérez P, Gianola D, Burgueño J, Araus JL, Makumbi D, Singh RP, Dreisigacker S, Yan J, Arief V. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*. 2010 Oct 1;186(2):713-24.
81. Juliana P, Poland J, Huerta-Espino J, Shrestha S, Crossa J, Crespo-Herrera L, Toledo FH, Govindan V, Mondal S, Kumar U, Bhavani S. Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nature genetics*. 2019 Oct;51(10):1530-9.
82. Kumar U, Kumar S, Prasad R, Röder MS, Kumar S, Chand R, Mishra VK, Joshi AK. Genetic Gain on Resistance to Spot Blotch of Wheat by Developing Lines with Near Immunity. *Crop Breeding, Genetics and Genomics*. 2019 Sep 6;1(2).
83. Juliana P, Singh RP, Singh PK, Poland JA, Bergstrom GC, Huerta-Espino J, Bhavani S, Crossa J, Sorrells ME. Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. *Theoretical and applied genetics*. 2018 Jul 1;131(7):1405-22.
84. Technow F, Messina CD, Totir LR, Cooper M. Integrating crop growth models with whole genome prediction through approximate Bayesian computation. *PloS one*. 2015 Jun 29;10(6):e0130855.
85. Crain J, Reynolds M, Poland J. Utilizing high-throughput phenotypic data for improved phenotypic selection of stress-adaptive traits in wheat. *Crop Science*. 2017;57(2):648-59.
86. Crossa J, de los Campos G, Maccaferri M, Tuberosa R, Burgueño J, Pérez-Rodríguez P. Extending the marker × environment interaction model for genomic-enabled prediction and genome-wide association analysis in durum wheat. *Crop Science*. 2016;56(5):2193-209.
87. Jarquín D, Lemes da Silva C, Gaynor RC, Poland J, Fritz A, Howard R, Battenfield S, Crossa J. Increasing genomic-enabled prediction accuracy by modeling genotype × environment interactions in Kansas wheat. *The plant genome*. 2017 Jul 1;10(2).
88. Saghai-Marouf MA, Soliman KM, Jorgensen RA, Allard RW. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences*. 1984 Dec 1;81(24):8014-8.
89. Dreisigacker S, Tiwari R, Sheoran S. Laboratory manual: ICAR-CIMMYT molecular breeding course in wheat. ICAR/BMZ/CIMMYT, Haryana, India, 2013;p 36
90. Poland JA, Brown PJ, Sorrells ME, Jannink JL. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS one*. 2012 Feb 28;7(2):e32253.
91. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007 Jun 22;23(19):2633-5.

92. Browning BL, Browning SR. Genotype imputation with millions of reference samples. *The American Journal of Human Genetics*. 2016 Jan 7;98(1):116-26.
93. Alvarado G, López M, Vargas M, Pacheco Á, Rodríguez F, Burgueño J, Crossa J. META-R (Multi Environment Trial Analysis with R for Windows) Version 6.03 CIMMYT Research Data & Software Repository Network: 2019. Available online: <https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10201>.
94. Wang J, Li H, Zhang L, Meng L. Users' Manual of QTL IciMapping Version 4.0, The Quantitative Genetics Group, Institute of Crop Science, Chinese Academy of Agricultural Sciences CAAS), Beijing 100081, China, and Genetic Resources Program, International Maize and Wheat Improvement Center CIMMYT), Apdo.
95. Team RC. R: A Language and Environment for Statistical Computing. 2019 Vienna: R Foundation for Statistical Computing.
96. Owens BF, Lipka AE, Magallanes-Lundback M, Tiede T, Diepenbrock CH, Kandianis CB, Kim E, Cepela J, Mateos-Hernandez M, Buell CR, Buckler ES. A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels. *Genetics*. 2014 Dec 1;198(4):1699-716.
97. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823. 2014 Jun 23.
98. Bischoff V, Selbig J, Scheible WR. Involvement of TBL/DUF231 proteins into cell wall biology. *Plant signaling & behavior*. 2010 Aug 1;5(8):1057-9.
99. Yang H, Li Y, Hua J. The C2 domain protein BAP1 negatively regulates defense responses in Arabidopsis. *The Plant Journal*. 2006 Oct;48(2):238-48..
100. Gunupuru LR, Arunachalam C, Malla KB, Kahla A, Perochon A, Jia J, Thapa G, Doohan FM. A wheat cytochrome P450 enhances both resistance to deoxynivalenol and grain yield. *PloS one*. 2018 Oct 12;13(10):e0204992.
101. Piesik D, Lemńczyk G, Skoczek A, Lamparski R, Bocianowski J, Kotwica K, Delaney KJ. Fusarium infection in maize: volatile induction of infected and neighboring uninfected plants has the potential to attract a pest cereal leaf beetle, *Oulema melanopus*. *Journal of plant physiology*. 2011 Sep 1;168(13):1534-42.
102. Zheng B, MacDonald TM, Sutinen S, Hurry V, Clarke AK. A nuclear-encoded ClpP subunit of the chloroplast ATP-dependent Clp protease is essential for early development in *Arabidopsis thaliana*. *Planta*. 2006 Oct 1;224(5):1103-15.
103. Gettins PG, Patston PA, Olson ST. *Serpins: structure, function and biology*. 1996;Austin: RG Landes Co
104. Antolín Llovera M, Petutsching EK, Ried MK, Lipka V, Nürnberger T, Robatzek S, Parniske M. Knowing your friends and foes—plant receptor like kinases as initiators of symbiosis or defence. *New Phytologist*. 2014 Dec 1;204(4):791-802.

105. Lu Y, Li Y, Zhang J, Xiao Y, Yue Y, Duan L, Zhang M, Li Z. Overexpression of Arabidopsis molybdenum cofactor sulfurase gene confers drought tolerance in maize (*Zea mays* L.). *PLoS one*. 2013 Jan 10;8(1):e52126.
106. Baniulis D, Hasan SS, Stofleth JT, Cramer WA. Mechanism of enhanced superoxide production in the cytochrome b₆ f complex of oxygenic photosynthesis. *Biochemistry*. 2013 Dec 6;52(50):8975-83.
107. Dilks T, Halsey K, De Vos RP, Hammond-Kosack KE, Brown NA. Non-canonical fungal G-protein coupled receptors promote *Fusarium* head blight on wheat. *PLoS pathogens*. 2019 Apr 1;15(4):e1007666.
108. Li J, Zhou F, Zhan D, Gao Q, Cui N, Li J, Iakhsiaeva E, Zwieb C, Lin B, Wong J. A novel histone H4 arginine 3 methylation-sensitive histone H4 binding activity and transcriptional regulatory function for signal recognition particle subunits SRP68 and SRP72. *Journal of Biological Chemistry*. 2012 Nov 23;287(48):40641-51.
109. Hu P, Liu J, Xu J, Zhou C, Cao S, Zhou W, Huang Z, Yuan S, Wang X, Xiao J, Zhang R. A malectin-like/leucine-rich repeat receptor protein kinase gene, RLK-V, regulates powdery mildew resistance in wheat. *Molecular plant pathology*. 2018 Dec;19(12):2561-74.
110. Favory JJ, Stec A, Gruber H, Rizzini L, Oravec A, Funk M, Albert A, Cloix C, Jenkins GI, Oakeley EJ, Seidlitz HK. Interaction of COP1 and UVR8 regulates UVB induced photomorphogenesis and stress acclimation in Arabidopsis. *The EMBO journal*. 2009 Mar 4;28(5):591-601.
111. Jenkins GI. Signal transduction in responses to UV-B radiation. *Annual review of plant biology*. 2009 Jun 2;60:407-31.
112. Chen D, Richardson T, Chai S, Lynne McIntyre C, Rae AL, Xue GP. Drought-up-regulated TaNAC69-1 is a transcriptional repressor of TaSHY2 and TaIAA7, and enhances root length and biomass in wheat. *Plant and Cell Physiology*. 2016 Jul 20;57(10):2076-90.
113. Collins NC, Thordal-Christensen H, Lipka V, Bau S, Kombrink E, Qiu JL, Hüchelhoven R, Stein M, Freialdenhoven A, Somerville SC, Schulze-Lefert P. SNARE-protein-mediated disease resistance at the plant cell wall. *Nature*. 2003 Oct;425(6961):973..
114. Yang KZ, Xia C, Liu XL, Dou XY, Wang W, Chen LQ, Zhang XQ, Xie LF, He L, Ma X, Ye D. A mutation in Thermosensitive Male Sterile 1, encoding a heat shock protein with DnaJ and PDI domains, leads to thermosensitive gametophytic male sterility in Arabidopsis. *The Plant Journal*. 2009 Mar;57(5):870-82.
115. Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *The Plant Cell*. 2004 Aug 1;16(8):2089-103.
116. Hwang IS, Choi DS, Kim NH, Kim DS, Hwang BK. The pepper cysteine/histidine-rich DC1 domain protein CaDC1 binds both RNA and DNA and is required for plant cell death and defense response. *New Phytologist*. 2014 Jan 1;201(2):518-30.
117. Li F, Asami T, Wu X, Tsang EW, Cutler AJ. A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant physiology*. 2007 Sep 1;145(1):87-97.

118. Dai X, Mashiguchi K, Chen Q, Kasahara H, Kamiya Y, Ojha S, DuBois J, Ballou D, Zhao Y. The biochemical mechanism of auxin biosynthesis by an Arabidopsis YUCCA flavin-containing monooxygenase. *Journal of Biological Chemistry*. 2013 Jan 18;288(3):1448-57.
119. Ergen NZ, Thimmapuram J, Bohnert HJ, Budak H. Transcriptome pathways unique to dehydration tolerant relatives of modern wheat. *Functional & integrative genomics*. 2009 Aug 1;9(3):377-96.
120. Mackey D, Holt III BF, Wiig A, Dangl JL. RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell*. 2002 Mar 22;108(6):743-54.
121. <https://www.uniprot.org/uniprot/A2X1N2>
122. Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature genetics*. 2007 May;39(5):623.
123. Liu TB, Chen GQ, Min H, Lin FC. MoFLP1, encoding a novel fungal fasciclin-like protein, is involved in conidiation and pathogenicity in *Magnaporthe oryzae*. *Journal of Zhejiang University Science B*. 2009 Jun 1;10(6):434-44.
124. Perochon A, Kahla A, Vranić M, Jia J, Malla KB, Craze M, Wallington E, Doohan FM. A wheat NAC interacts with an orphan protein and enhances resistance to *Fusarium* Head Blight disease. *Plant biotechnology journal*. 2019 Mar 1.
125. Lionetti V, Raiola A, Camardella L, Giovane A, Obel N, Pauly M, Favaron F, Cervone F, Bellincampi D. Overexpression of pectin methylesterase inhibitors in Arabidopsis restricts fungal infection by *Botrytis cinerea*. *Plant physiology*. 2007 Apr 1;143(4):1871-80.
126. Dmochowska Boguta M, Nadolska Orczyk A, Orczyk W. Roles of peroxidases and NADPH oxidases in the oxidative response of wheat (*Triticum aestivum*) to brown rust (*Puccinia triticina*) infection. *Plant Pathology*. 2013 Oct;62(5):993-1002.
127. Sappl PG, Carroll AJ, Clifton R, Lister R, Whelan J, Harvey Millar A, Singh KB. The Arabidopsis glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. *The Plant Journal*. 2009 Apr;58(1):53-68.
128. Beffa RS, Neuhaus JM, Meins F. Physiological compensation in antisense transformants: specific induction of an "ersatz" glucan endo-1, 3-beta-glucosidase in plants infected with necrotizing viruses. *Proceedings of the National Academy of Sciences*. 1993 Oct 1;90(19):8792-6.
129. Gutsch A, Keunen E, Guerriero G, Renaut J, Cuypers A, Hausman JF, Sergeant K. Long-term cadmium exposure influences the abundance of proteins that impact the cell wall structure in *Medicago sativa* stems. *Plant Biology*. 2018 Nov;20(6):1023-35.
130. Millyard L, Lee J, Zhang C, Yates G, Sadanandom A. The ubiquitin conjugating enzyme, TaU4 regulates wheat defence against the phytopathogen *Zymoseptoria tritici*. *Scientific reports*. 2016 Oct 19;6:35683.
131. Uni Prot KB. Swiss Prot. 2017 Mar 15

132. Morkunas I, Ratajczak L. The role of sugar signaling in plant defense responses against fungal pathogens. *Acta Physiologiae Plantarum*. 2014 Jul 1;36(7):1607-19.
133. Shukla V, Kaur M, Aggarwal S, Bhati KK, Kaur J, Mantri S, Pandey AK. Tissue specific transcript profiling of wheat phosphate transporter genes and its association with phosphate allocation in grains. *Scientific reports*. 2016 Dec 20;6:39293.
134. Cui XY, Du YT, Fu JD, Yu TF, Wang CT, Chen M, Chen J, Ma YZ, Xu ZS. Wheat CBL-interacting protein kinase 23 positively regulates drought stress and ABA responses. *BMC plant biology*. 2018 Dec;18(1):93.
135. Mock HP, Keetman U, Kruse E, Rank B, Grimm B. Defense responses to tetrapyrrole-induced oxidative stress in transgenic plants with reduced uroporphyrinogen decarboxylase or coproporphyrinogen oxidase activity. *Plant physiology*. 1998 Jan 1;116(1):107-16.
136. Strygina KV, Khlestkina EK. Myc-like transcriptional factors in wheat: structural and functional organization of the subfamily I members. *BMC plant biology*. 2019 Feb;19(1):50.
137. Hückelhoven R. Cell wall-associated mechanisms of disease resistance and susceptibility. *Annu. Rev. Phytopathol.*. 2007 Sep 8;45:101-27.
138. Yao X, Xiong W, Ye T, Wu Y. Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in *Arabidopsis*. *Journal of experimental botany*. 2012 Jan 20;63(7):2579-93.
139. La Camera S, Geoffroy P, Samaha H, Ndiaye A, Rahim G, Legrand M, Heitz T. A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in *Arabidopsis*. *The Plant Journal*. 2005 Dec;44(5):810-25.
140. Zhu X, Wang Y, Su Z, Lv L, Zhang Z. Silencing of the wheat protein phosphatase 2A catalytic subunit TaPP2Ac enhances host resistance to the necrotrophic pathogen *Rhizoctonia cerealis*. *Frontiers in Plant Science*. 2018 Oct 31;9:1437.

Tables

Table 1: Range, mean values, standard deviation and sample variance of 141 advanced lines evaluated for spot blotch disease resistance at growth stage 77 on Zadoks Scale in four environments

Statistic	Env1	Env2	Env3	Env4
Mean	21.67	8.90	15.10	31.23
Standard Error	0.88	0.71	0.39	0.72
Standard Deviation	14.81	11.96	6.55	12.03
Sample Variance	219.45	143.09	42.88	144.83
Range	0-70	0-50	0-43	11-66
Confidence Level (95.0%)	1.73648152	1.402192864	0.767555673	1.410680498

Table 2: Analysis of variance of 141 lines evaluated for spot blotch disease resistance in four environments based on BLUP of disease severity recorded at growth stage 77 on Zadoks scale

Statistic	Env1	Env2	Env3	Env4	Combined
Heritability	0.76	0.71	0.50	0.80	0.55
LSD	16.49	14.16	14.41	8.74	11.04
CV	28.21	30.06	29.73	22.41	28.53
Replications	2	2	2	2	2
Residual variance	184.28	146.56	222.91	48.98	151.63
Gen variance	284.48	174.91	113.18	96.25	68.80
Gen significance	1.11022E-15	2.6390E-13	4.73255E-05	0	2.82891E-08
Gen × Env variance	-	-	-	-	96.00
Gen × Env significance	-	-	-	-	3.69898E-14
Env variance	-	-	-	-	104.36
Env significance	-	-	-	-	0.0115412318

Table 3. List of significant SNPs with the corresponding proteins and possible function elucidated based on the gene annotation using wheat reference sequence (RefSeq V1.0) database

SNP markers	GeneID	Gene annotation	Possible Function/Description	References
ENV1 (PUSA17)				
S1A_9565863	TraesCS1A01G018700	Trichome birefringence	Cellulose biosynthesis	[98]
	TraesCS1A01G018800	Transmembrane protein	Regulates fungal development and pathogenicity via MAPK module FgSte50-Ste11-Ste7 in <i>F. graminearum</i>	[69]
S1D_6541259	TraesCS1D01G012500	C2 calcium/lipid-binding and GRAM domain protein	C2 domain protein BAP1 negatively regulates defense responses in <i>Arabidopsis</i>	[99]
S1D_6715588	TraesCS1D01G012800	Serine/threonine-protein kinase	Confers powdery mildew resistance in wheat	[66]
	TraesCS1D01G012900			
S2B_13751999	TraesCS2B01G030100	Disease resistance protein RPP13,	Play important roles in the resistance of various plant diseases including the downy mildew	[31]
	TraesCS2B01G030200	lectin-receptor kinase	Lectin receptor kinases are involved in plant immunity	[32]
S2B_13761590	TraesCS2B01G030200	lectin-receptor kinase	Lectin receptor kinases are involved in plant immunity	[32]
S2B_13814702	TraesCS2B01G030500	Cytochrome P450 family protein	Enhances both resistance to deoxynivalenol and grain yield	[100]
S2B_14261851	TraesCS2B01G030900	Myrcene synthase	Interference and disruption of membranes in maize infected by <i>Fusarium spp</i>	[101]
	TraesCS2B01G031000			
S2B_14809954	TraesCS2B01G031700	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
	TraesCS2B01G031800	ATP-dependent Clp protease adapter protein	Essential for early development in <i>Arabidopsis</i>	[102]
S2B_14963432	TraesCS2B01G031900	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
	TraesCS2B01G032000			
S2B_15129248	TraesCS2B01G032100	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[56]
	TraesCS2B01G032200	Zinc finger family protein (GRF)	Involved in plant disease resistance	
S3B_763230831	TraesCS3B01G519800	Transmembrane protein	Transmembrane protein FgSho1 regulates fungal development and pathogenicity via MAPK module Ste50-Ste11-Ste7 in <i>F. graminearum</i>	[69]
	TraesCS3B01G519900	General regulatory factor 1	-	-
S3B_763236179	TraesCS3B01G520000	Serpin family protein	Participate in the regulation of complex proteolytic systems	[103]
S3B_763267753	TraesCS3B01G520100	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	Enhances resistance to deoxynivalenol and increase grain yield	[100]
	TraesCS3B01G520200	Serpin family protein	Participate in the regulation of complex proteolytic systems	[103]
S3B_764173978	TraesCS3B01G520800	-	-	-
	TraesCS3B01G520900	F-box family protein	Negative regulator of the defense response	[58]
S3B_764747435	TraesCS3B01G521400	Glycerol-3-phosphate dehydrogenase [NAD(+)]	Contributes to systemic acquired resistance against <i>P. striiformis f. sp. tritici</i>	[59]
	TraesCS3B01G521500	Zinc finger family protein	Plant disease resistance	[56]
S3B_764804887	TraesCS3B01G521600	Zinc finger family protein	Plant disease resistance	[56]
	TraesCS3B01G521700	Tyrosyl-DNA phosphodiesterase	-	-
S5A_593739698	TraesCS5A01G400900	Kinase family protein	As initiators of symbiosis or defense	[104]
	TraesCS5A01G401000	Keratin, type II cuticular Hb1	-	-
S5A_595158840	TraesCS5A01G402500	Molybdenum cofactor sulfurase	Overexpression of <i>Arabidopsis</i> Molybdenum Cofactor Sulfurase gene confers drought tolerance in maize	[105]
	TraesCS5A01G402600	Cytochrome b6-f complex subunit 7	Oxygenic photosynthesis	[106]

S5A_595373332	TraesCS5A01G402700	Transmembrane protein	Regulates fungal development and pathogenicity via MAPK module FgSte50-Ste11-Ste7 in <i>F. graminearum</i>	[69]
S5A_595393566	TraesCS5A01G402800	Myb family transcription factor-like protein	Mediates host resistance to <i>Bipolaris sorokiniana</i> in wheat	[57]
S5A_636959783	TraesCS5A01G457100	Guanylate-binding family protein	Non-canonical fungal G-protein coupled receptors promote Fusarium head blight on wheat	[107]
	TraesCS5A01G457200	carboxyl-terminal peptidase	-	-
S5B_233586644	TraesCS5B01G127900	Hippocampus abundant transcript-like protein 1	-	-
	TraesCS5B01G128000	Signal recognition particle subunit SRP68	Crucial role in targeting secretory proteins to the rough endoplasmic reticulum membrane	[108]
S6A_29496364	TraesCS6A01G056000	F-box family protein	Negative regulator of the defense response	[58]
S6A_33023854	TraesCS6A01G061900	Kinase family protein	As initiators of symbiosis or defense	[104]
	TraesCS6A01G062000	F-box family protein	Negative regulator of the defense response	[58]
S6A_33299007	TraesCS6A01G062300	DNA-directed RNA polymerase II subunit rpb4	Pol II alone is capable of RNA transcript elongation and of proof reading	-
	TraesCS6A01G062400	Pigment epithelium-derived factor	-	-
S7B_1020705	TraesCS7B01G002400	Leucine-rich repeat receptor-like protein kinase	Pathogen-Responsive Leucine Rich Receptor Like Kinase Contributes to <i>Fusarium</i> Resistance in Cereals and regulates powdery mildew resistance in wheat	[53, 109]
S7B_1261577	TraesCS7B01G003000	COP1-interacting-like protein	COP1 and HY5 also play essential roles during UV-B signaling	[110, 111]
ENV2 (PUSA18)				
S1B_647195634	TraesCS1B01G424000	Pentatricopeptide repeat-containing protein	Involved in plant disease resistance	[65]
	TraesCS1B01G424100	Homeobox protein	Associated with Reaction to Stripe Rust and Powdery Mildew	[72]
S2A_708482943	TraesCS2A01G462700	Receptor-like kinase	as initiators of symbiosis or defense	[104]
	TraesCS2A01G462800	Glycosyltransferase	Involved in plant disease resistance	[55]
S2A_709173869	TraesCS2A01G463300	Glycosyltransferase	Involved in plant disease resistance	[55]
S2B_90662917	TraesCS2B01G123200	Transcription repressor ofp17	Transcriptional Repressor of TaSHY2 and TaIAA7, Enhances Root Length and Biomass in Wheat	[112]
	TraesCS2B01G123300	-	-	-
S3D_610628298	TraesCS3D01G537500	Zinc finger protein (Dof)	Zinc Finger Proteins Involved in Plant Disease Resistance	[56]
S5B_400097303	TraesCS5B01G224500	Golgi SNAP receptor complex member 1	Involved in plant disease resistance through regulation of Plant Cell Death	[113]
S5B_663822910	TraesCS5B01G496700	Chaperone protein dnaJ	leads to thermosensitive gametophytic male sterility in Arabidopsis	[114]
	TraesCS5B01G496800	Chaperone protein dnaJ	leads to thermosensitive gametophytic male sterility in Arabidopsis	
S7A_709769485	TraesCS7A01G530700	Pentatricopeptide repeat-containing protein	Essential Role in Organelle Biogenesis	[115]
	TraesCS7A01G530800	Cysteine/histidine-rich C1 domain-containing protein	Pepper cysteine/histidine-rich DC1 domain protein gene, CaDC1, positively regulates plant defense during microbial infection	[116]
S7D_181974079	TraesCS7D01G221000	Hydroxysteroid	Involved in regulating plant growth and development	[117]
	TraesCS7D01G221100	dehydrogenase		
ENV3 (PUSA19)				
S2A_31851904	TraesCS2A01G071500	Flavin-containing	Mediate two-step auxin biosynthesis pathway in <i>Arabidopsis</i> .	[118]

	TraesCS2A01G071600	monoxygenase		
S2B_504717	TraesCS2B01G001100	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
	TraesCS2B01G001200	Chaperone protein DnaJ	Leads to thermo-sensitive gametophytic male sterility in <i>Arabidopsis</i>	Yang et al. 2009 114
S2B_319090	TraesCS2B01G000700	3'-N-debenzoyl-2'-deoxytaxol N-benzoyltransferase	-	-
S2B_78065	TraesCS2B01G000400	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
	TraesCS2B01G000500	Peptide transporter	-	-
S2B_8311062	TraesCS2B01G018100	Terpene synthase	A potential role of terpenes in drought tolerance	[119]
	TraesCS2B01G018200	NBS-LRR disease resistance protein-like protein	Disease resistance protein	[54, 55]
S2B_6253562	TraesCS2B01G012400	Avr9/Cf-9 rapidly elicited protein	Early signaling events in the Avr9/Cf-9-dependent plant defence response	[61]
S2B_90662917	TraesCS2B01G123200	Transcription repressor ofp17	Transcriptional repressor of TaSHY2 and TaIAA7, enhances root length and biomass in wheat	[112]
	TraesCS2B01G123300	-	-	-
S2B_89540368	TraesCS2B01G122000	Mitochondrial transcription termination factor-like	-	-
	TraesCS2B01G122100	Expansin protein	-	-
S2B_748595375	TraesCS2B01G552600	Disease resistance protein RPM1	Disease resistance protein	[120]
	TraesCS2B01G552700	Signal peptide, CUB and EGF-like domain-containing protein 3	-	-
S3A_67348475	TraesCS3A01G104000	helicase with zinc finger protein		
	TraesCS3A01G104100	Two-component response regulator ORR24	Involved in His-to-Asp phosphorelay signal transduction system	[121]
S3A_66513067	TraesCS3A01G103000	RING/U-box superfamily protein	Associated with the control of grain size	[122]
	TraesCS3A01G103100	Fasciclin-like arabinogalactan protein	Involved in conidiation and pathogenicity in <i>Magnaporthe oryzae</i>	[123]
S3A_67065083	TraesCS3A01G103400	Di-glucose binding protein with Kinesin motor domain-containing protein	-	-
	TraesCS3A01G103500	Transcription factor-like protein	1R-MYB Transcription Factor, plays an important role in disease resistance against stripe rust fungus and ear heading in wheat	[45]
S3A_72072539	TraesCS3A01G107300	Zinc finger family protein (GRF)	Involved in plant disease resistance	[56]
	TraesCS3A01G107400	NAC domain protein	Wheat NAC interacts with an orphan protein and enhances resistance to FHB disease	[124]
S5B_673038030	TraesCS5B01G507700	Protein plant cadmium	-	-

		resistance		
	TraesCS5B01G507800	Invertase inhibitor	Overexpression of Pectin Methylesterase Inhibitors in <i>Arabidopsis</i> restricts fungal infection by <i>Botrytis cinerea</i>	[125]
S6A_23959761	TraesCS6A01G047300	Peroxidase	Divergent role in different plant-pathogen systems	[126]
	TraesCS6A01G047400			
S7B_240325318	TraesCS7B01G169400	Glutathione S-transferase T3	Involved in complex stress regulation in <i>Arabidopsis</i>	[127]
	TraesCS7B01G169500	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
S7D_49641314	TraesCS7D01G082700	RING/U-box superfamily protein	Associated with the control of grain size	[122]
	TraesCS7D01G082800	Polygalacturonase-1 non-catalytic beta subunit	Implicated in the defense against fungal pathogens in Plants through cell wall modification, degradation, carbohydrate metabolic processes and as response to stress	[128, 129]
S7D_47848266	TraesCS7D01G081100	Telomerase activating protein	-	-
	TraesCS7D01G081200	Ubiquitin-conjugating enzyme	Regulates wheat defence against the phytopathogen <i>Zymoseptoria tritici</i>	[130]
ENV4 (BARI17)				
S1B_646895451	TraesCS1B01G423900	CXC domain-containing protein	Role in reproductive tissues development	[131]
	TraesCS1B01G424000	Pentatricopeptide repeat-containing protein	Involved in plant disease resistance	[65]
S2B_28592818	TraesCS2B01G058800	Sugar transporter	Role in plant defense responses against fungal pathogens	[132]
	TraesCS2B01G058900	NBS-LRR disease resistance protein	Disease resistance protein	[54, 55]
S2B_533178164	TraesCS2B01G373800	little nuclei4	-	-
	TraesCS2B01G373900	Cytochrome P450	Enhances resistance to deoxynivalenol and increase grain yield	[100]
S2B_699219601	TraesCS2B01G505200	Disease resistance protein-like	Important roles in the resistance of various plant diseases including downy mildew	[31]
S4A_710830493	TraesCS4A01G443100	Sodium-dependent phosphate transporter	Importance in the filial tissues during grain filling	[133]
	TraesCS4A01G443200	ATP-dependent Clp protease ATP-binding subunit	Essential for early development in <i>Arabidopsis</i>	[102]
S4B_368023618	TraesCS4B01G168200	CBL-interacting Serine/Threonine-kinase	Regulates drought stress and ABA responses	[134]
	TraesCS4B01G168300	Uroporphyrinogen decarboxylase	Defense responses	[135]
S5B_681773490	TraesCS5B01G518800	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	Regulate growth and development as well as response to various stresses	[136]
	TraesCS5B01G518900	Hydroxyproline-rich glycoprotein family protein	Provides added resistance against pathogen-derived cell wall-degrading enzymes	[137]
S5B_683352145	TraesCS5B01G521300	NBS-LRR disease resistance protein family-1	Disease resistance protein	[54, 55]
	TraesCS5B01G521400	-	-	-
S5B_683514734	TraesCS5B01G521400	-	-	-
	TraesCS5B01G521500	Kinase family protein	As initiators of symbiosis or defense	[104]

S5B_684060726	TraesCS5B01G522300	F-box family protein	Negative regulator of the defense response	[58]
	TraesCS5B01G522400	Glycosyltransferase	Involved in plant disease resistance	[55]
S5B_684129768	TraesCS5B01G522500	ATP-dependent zinc metalloprotease FTSH protein	-	-
	TraesCS5B01G522600	Eukaryotic aspartyl protease family protein	Overexpression of the ASPG1 gene confers drought avoidance in <i>Arabidopsis</i>	[138]
S5B_684230407	TraesCS5B01G522900	Protein detoxification	-	-
S5B_684610070	TraesCS5B01G523800	ATP-dependent zinc metalloprotease	Regulates hydrolase, metalloprotease and protease activities - cell division (peptidase/protease)	-
S5B_74329592	TraesCS5B01G066100	Invertase/pectin methylesterase inhibitor family protein	Overexpression of Pectin Methylesterase inhibitors in <i>Arabidopsis</i> restricts fungal infection by <i>Botrytis cinerea</i>	[125]
	TraesCS5B01G066200	PME/invertase inhibitor-like protein		
S6B_719904092	TraesCS6B01G472300	RING/U-box superfamily protein	Associated with the control of grain size	[122]
	TraesCS6B01G472400	Integral membrane metal-binding family protein	-	-
S7A_692811400	TraesCS7A01G504700	Patatin	Pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in <i>Arabidopsis</i>	[139]
	TraesCS7A01G504800	Protein phosphatase 2c	Function of the plant PP2Ac genes in plant immune responses	[140]
S7D_37950527	TraesCS7D01G067000	-	-	-
	TraesCS7D01G067100	Phosphate-responsive family protein	Function of the plant PP2Ac genes in plant immune responses	[140]

Figures

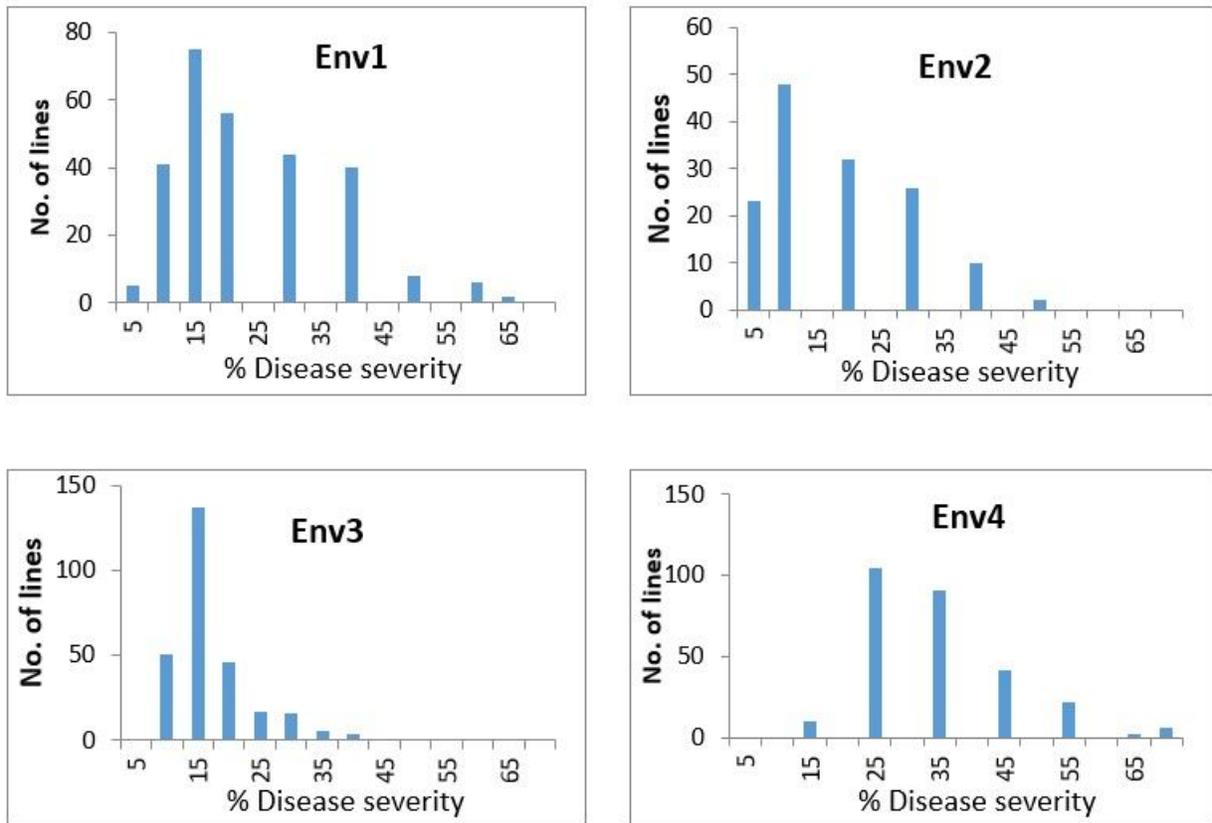


Figure 1

Distribution of lines for spot blotch resistance and susceptibility evaluated in four environments

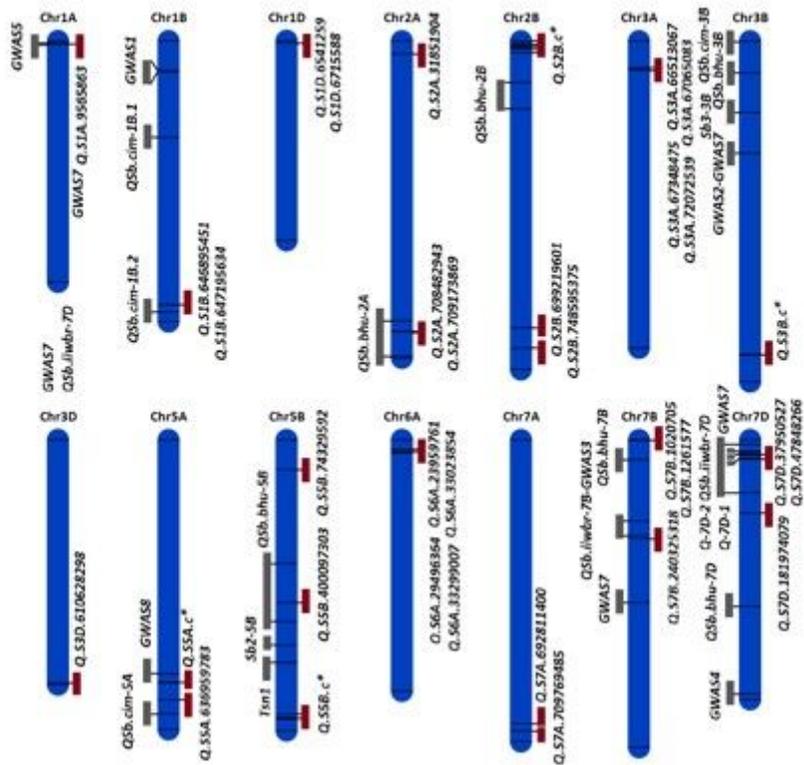


Figure 2

Physical map of QTLs mapped along with already known/mapped loci. The gray and red rectangles represent loci from literature and mapped during this study respectively. The loci marked with (c*) represent clusters of loci mapped at a particular region. Q.S2B.c: Q.S2B.319090, Q.S2B.6253562, Q.S2B.8311062, Q.S2B.13751999, Q.S2B.13761590, Q.S2B.13814702, Q.S2B.14261851, Q.S2B.14809954, Q.S2B.14963432, Q.S2B.15129248, Q.S2B.19097998, Q.S2B.28592818; Q.S3B.c: Q.S3B.763230831, Q.S3B.763236179, Q.S3B.763267753, Q.S3B.764173978, Q.S3B.764747435, Q.S3B.764804887; Q.S5A.c: Q.S5A.593739698, Q.S5A.595158840, Q.S5A.595373332, Q.S5A.595393566, Q.S5A.636959783; and Q.S5B.c: Q.S5B.673038030, Q.S5B.681773490, Q.S5B.683352145, Q.S5B.683514734, Q.S5B.684060726, Q.S5B.684129768, Q.S5B.684230407, Q.S5B.684610070

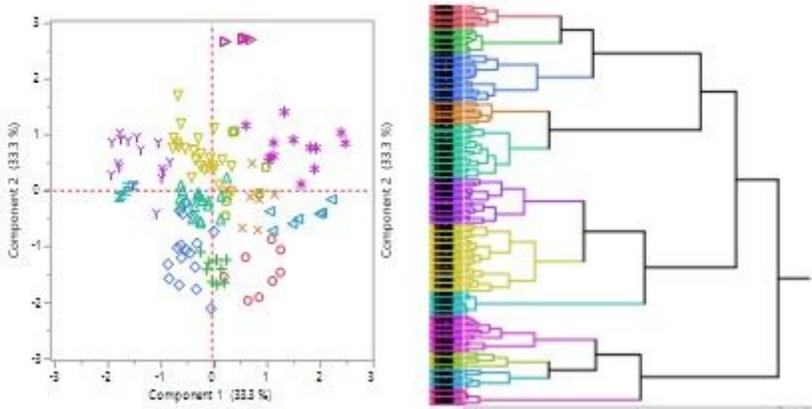


Figure 3

Population structure: PCA and Dendrogram

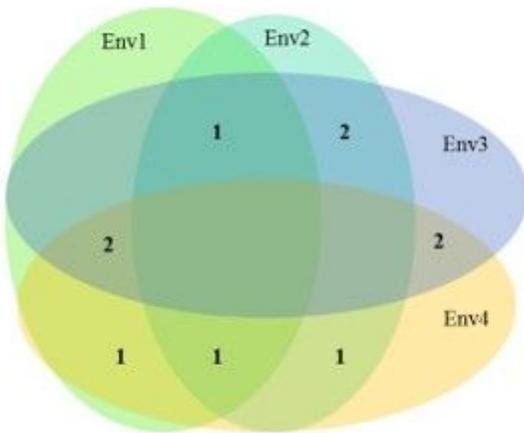


Figure 4

Venn diagram based on the common protein synthesized by the same genes associated with different SNPs over years (The numbers represent the total proteins common between the environments)

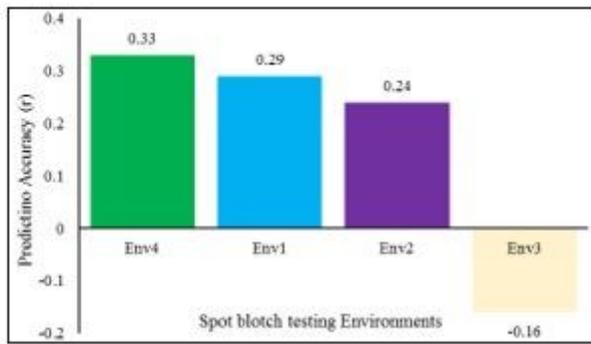


Figure 5

Genomic prediction for spot blotch in four different environments using rrBLUP

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [7TableS3Weather.docx](#)
- [6TableS2Commonproteins.xlsx](#)
- [5TableS1SignificantSNPs.xlsx](#)