

Genome-Wide Association Mapping of Major QTLs for Resistance to Bacterial Leaf- Streak Disease (*Xanthomonas translucence* pv. *undulosa*) in Elite Hard Winter Wheat Germplasm

Muhammad Ahmad

South Dakota State University

Jeffrey D. Boehm

University of Nebraska-Lincoln East Campus, USDA-ARS

Paul St. Amand

USDA-ARS

Amy Bernardo

USDA-ARS

Katherine Jordan

USDA-ARS

Guihua Bai

USDA-ARS

Meriem Aoun

Oklahoma State University

Hamza Ashfaq

University of Florida, IFAS

Karl D. Glover

South Dakota State University

Shyam Solanki

South Dakota State University

Gazala Ameen

gazala.ameen@sdstate.edu

South Dakota State University

Research Article

Keywords: Bacterial Leaf Streak Disease, wheat (*Triticum aestivum*), Genome-wide association study, Resistant genes

Posted Date: March 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9031856/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on September 12th, 2026. See the published version at <https://doi.org/10.1186/s12864-026-13325-2>.

Abstract

Background: Bacterial leaf streak (BLS), caused by *Xanthomonas translucens* pv. *undulosa* (*Xtu*), is an increasing threat to wheat (*Triticum aestivum* L.) production in the Northern Great Plains of the United States, where severe epidemics can cause up to 60% yield losses in susceptible genotypes [1]. Despite its growing economic impact, genetic resistance in wheat remains poorly characterized, and only limited resistant germplasm has been identified. For that purpose, this study aimed to evaluate a diverse Regional Germplasm Observation Nursery (RGON) panel composed of elite contemporary winter wheat lines for BLS resistance and to identify genomic regions associated with resistance using genome-wide association studies (GWAS).

Results: A total of 412 elite hard winter wheat lines were phenotyped for BLS response under controlled greenhouse conditions using a 1-9 disease severity scale. A broad range of phenotypic responses was observed, although most accessions exhibited moderate susceptibility. Multiplex Restriction Amplicon Sequencing (MRASeq) produced 15,368 high-quality SNPs after quality filtering. Population structure analyses revealed five genetically distinct subpopulations. GWAS using BLINK and FarmCPU models identified seven significant marker trait associations (MTAs) on chromosomes 1B, 2B, 3B, 3D, 4A, 4B, and 6B, respectively, explaining a range of 0.04% to 16.6 % of the phenotypic variance. Out of the seven, four are major QTL explaining >10% phenotypic variance. Chromosome 3B showed strong co-occurrence with previously reported resistance loci, while several loci on 2B, 4A, 4B, 1B, 6B, and 3D represent putative novel regions for BLS resistance. Seventeen highly resistant lines consistently carried favorable alleles across major loci on chromosomes 3B, 4A, 4B, and 6B.

Conclusions: This study elucidates the genetic architecture of BLS resistance in hard winter wheat adapted to the Northern Great Plains of the US and identifies both known and novel resistance loci. The resistant germplasm and SNP markers reported herein will provide valuable resources for marker-assisted breeding to accelerate the development of BLS-resistant winter wheat cultivars.

Background

Wheat (*Triticum aestivum* L.) holds significant global importance as a cereal crop, with hard winter wheat alone contributing to 40–50% of total wheat production in the USA (United States Department of Agriculture, 2023). However, hard winter wheat faces numerous biotic and abiotic challenges that limit yield and seed quality (Pandey et al., 2017). One emerging challenge is bacterial leaf streak (BLS) disease, which affects many small grain cereals including wheat, barley, rye, oat, and triticale. For wheat, BLS disease infections are caused by *Xanthomonas translucens* pv. *translucens* (*Xtt*) (Ritzinger et al., 2023), whereas pv. *undulosa* (*Xtu*) infects both wheat and barley (Sapkota et al., 2020). Common symptoms of BLS include water-soaked, translucent streaks that progress into chlorosis and then form necrotic lesions on leaves, sometimes resulting in the entire plant displaying a blight appearance (Heiden et al., 2022). BLS infections on grain tissue and heads can lead to kernel darkening, known commonly as ‘black chaff’. In the Northern Great Plains, especially in North Dakota, South Dakota and Minnesota, BLS incidence has notably increased over the past decade (Hong et al., 2023; Kandel et al., 2012; Ramakrishnan et al., 2019), with yield losses on susceptible varieties reaching up to 60% when infections are severe (Friskop et al., 2023). On average, BLS and black chaff incidences on grain inspections have resulted in the loss of approximately 14.5 million US wheat bushels, amounting to \$78.5 million annually (Crop Protection Network, 2023). In 2018 and 2020 alone, losses to BLS totaled approximately 24 and 23 million bushels, resulting in revenue losses of \$121 million and \$119 million, respectively (Crop Protection Network, 2023). Recent studies have explored combating BLS disease using chemicals such as silicon (Silva et al., 2010) and even copper-containing compounds (Duveiller et al., 1997; Shohanipour et al., 2025). However, the best, most cost-effective solution remains developing varieties with genetic resistance to *X. translucens*. Unfortunately, the predominant wheat varieties grown in the Great Plains region have shown moderate to high susceptibility to BLS disease in recent years (Ramakrishnan et al., 2019).

Previous studies have extensively analyzed various types of wheat germplasm, including cultivars, breeding lines, and landraces, to assess their resistance or susceptibility to bacterial leaf streak (BLS) across diverse environmental conditions. However, no single genotype has shown significant resistance or provided immunity to the *Xtu* pathogen (Kandel et al., 2012; Ramakrishnan et al., 2019; Tillman et al., 1996). Duveiller et al. (1992) identified five genetic loci (*Bls1*, *Bls2*, *Bls3*, *Bls4*, and *Bls5*) within an F₃ population resulting from crossing five distinct parental lines, revealing partial resistance against BLS disease. Adhikari et al. (2012) reported a mere 8.3% prevalence of resistance among 605 winter wheat accessions obtained from the USDA National Small Grain Collection (NSGC). Ramakrishnan et al. (2019) evaluated 299 hard winter wheat lines and revealed ten resistant genotypes with five regions regulating resistance, explaining 42% of the variation. While numerous studies have investigated germplasm and identified resistant sources, few have focused on unraveling the genetic architecture underlying resistance to bacterial leaf streak (BLS) in spring wheat (Adhikari et al., 2012; Kandel et al., 2012; Wen et al., 2018). Adhikari et al. (2012) analyzed 566 spring wheat germplasms, revealing quantitative trait loci (QTLs) associated with BLS resistance on chromosomes 1A, 4A, 4B, and 6A, explaining phenotypic variance that ranged from 1.4 to 2.6%. Kandel et al. (2015a) identified QTLs on chromosomes 1B, 2A, 3B, and 6A within a diverse spring wheat population, with notable SSR markers elucidating phenotypic variance spanning up to 23%. Wen et al. (2018) discovered the presence of a BLS resistance gene in Triticale on chromosome 5R. Although BLS poses a significant challenge to rice cultivation, numerous genes and quantitative trait loci (QTLs) associated with resistance to BLS have been identified in the rice genome. (HE et al., 2012; Tang et al., 2000).

Quantitative traits, such as Bacterial Leaf Streak (BLS), are often genetically mapped using linkage analysis or association mapping (AM) methods. For the latter, AM utilizes historical recombination zones within sets of generally unrelated lines collected from natural collections, breeding populations, or regional varieties (Zhu et al., 2008). The emergence of high-throughput marker technologies has facilitated genome-wide association studies (GWAS) as an effective AM strategy for identifying Quantitative Trait Nucleotides (QTNs) and novel alleles, particularly in genetically diverse germplasm or wild relatives, rather than parental populations (Begum et al., 2015; Zhao et al., 2011). This approach eliminates the need for prior information on marker-trait associations and enables the convenient identification of multiple loci conditioning traits of interest (Bordes et al., 2011). Various robust statistical methods and tools have been developed to enhance our understanding of complex traits in animal and plant genetic systems. These include mixed linear models (Yu et al., 2006; Zhang et al., 2010), Bayesian clustering (Pritchard & Rosenberg, 1999; Pritchard et al., 2000), and principal component analysis (PCA) (Price et al., 2006). These methods have been extensively applied in numerous studies to characterize resistance to wheat diseases such as stem rust, the solid stem trait for sawfly resistance, leaf rust, spot blotch, and also in other major crops like soybean, rice and maize (Letta et al., 2014; Muleta et al., 2017; Raboin et al., 2016; Xiao et al., 2017). Hence, GWAS offers the potential for identifying genetic loci governing BLS resistance in winter wheat, which would facilitate the improvement of wheat varieties through the strategic integration of beneficial alleles using marker-assisted pyramiding and backcross breeding.

The objectives of the present study were to: (1) assess BLS resistance or susceptibility reactions among winter wheat accessions and breeding lines from the Regional Germplasm Observation Nursery (RGON) collected, and (2) use that same set of breeding lines to conduct GWAS to identify genomic regions associated with BLS resistance; and (3) to develop new SNP markers to aid breeders with marker-assisted selection of BLS resistance.

Methods

Bacterial Leaf Streak Strain and Plant Materials

The *Xtu* bacterial strain AL-1029 (*Xanthomonas translucens* pv. *undulosa*) (GenBank: PQ524066) used in this study was isolated from a wheat field in South Dakota during the summer 2023 growing season (Ahmad et al., 2024). A total of 412 hard winter wheat lines selected from the Regional Germplasm Observation Nursery (RGON) 2023 were evaluated for bacterial leaf streak (BLS) resistance or susceptibility. These lines were obtained from the United States Department of

Agriculture (USDA) Wheat, Sorghum and Forage Research Unit in Lincoln, Nebraska, which coordinates the Hard Winter Wheat Regional Performance Nurseries (HWWRPN). The HWWRPN aims to enhance the development of hard winter wheat varieties suited to diverse U.S. Agro-climatic regions through collaborative efforts between USDA, universities, research institutions, and private industry throughout the Northern and Southern Great Plains (2023 RGON 120823 1350.xlsx). The RGON wheat lines that were screened herein represent elite public and private breeding entries from multiple organizations representing nine U.S. states where majority of winter wheat is cultivated every year (Fig. 1).

Experimental Design and Phenotyping

The disease screening experiments were conducted under controlled greenhouse conditions over two consecutive years at the SDSU northwest greenhouse facility. Seeds were planted in 9-inch cones (two seeds per cone) following a completely randomized design (CRD) with three replications for a 16/8 hour light/dark condition. After 15 days of germination, six plants per genotype were inoculated with the bacterial strain *Xtu-1029* using a foliar spray method (Ahmad et al., 2025). Following the bacterial inoculation, plants were incubated in a misting chamber at 100% relative humidity for two days and subsequently maintained at 80% humidity for seven days at 23 to 25°C. Disease severity was recorded seven days post-inoculation using a 1–9 scale, where 1 indicates a highly resistant reaction and 9 indicates a highly susceptible reaction based on observed symptoms (Ahmad et al., 2025). The disease ratings were classified as highly resistant (scores 1–2), moderately resistant (scores 3–5), moderately susceptible (scores 6–7), and highly susceptible (scores 8–9). For further analysis, the mean of the three replicates was calculated for each line (Gomez & Gomez, 1984). Analysis of variance (ANOVA) on the phenotype scores was performed using R (version 4.5.2) to assess differences among lines. Homogeneity of variance among replicates was confirmed prior to analysis (**Supplementary Table 1**).

Genotyping

Genotyping was conducted using Multiplex Restriction Amplicon Sequencing (MRASeq), following the protocol described by Bernardo et al. (2020) at the Hard Winter Wheat Genetics Research Unit, USDA-ARS, Manhattan, KS, USA, to generate high-quality DNA sequence data. The SNP calling was carried out following the procedures described in TASSEL v5.0 (Bradbury et al., 2007). SNP markers with 10% missing data, minor allele frequency (MAF) < 0.05, and unmapped loci were filtered out. Missing genotypes were imputed using BEAGLE v4.1 (Browning & Browning, 2007). After filtering, a total of 15,368 high-quality SNPs were retained for downstream analyses. This stringent filtering procedure utilized herein ensured that marker–trait associations were accurate and reliable by minimizing false positives.

Genetic Diversity Analysis and Linkage Disequilibrium

Genetic relatedness among genotypes was estimated using the identity-by-state (IBS) distance matrix in TASSEL to calculate the proportion of shared alleles among all pairs of individuals. A neighbor-joining tree was constructed using DarWin software (Perrier, 2006) to visualize phylogenetic relationships. Principal component analysis (PCA) was performed in PLINK with the *pca* function, generating eigenvectors and eigenvalues for downstream visualization (Price et al., 2006). The first several principal components were plotted in R 4.4.2 using ggplot2 and tidyverse packages to visualize clustering patterns among accessions. Population structure was analyzed using ADMIXTURE v1.3.0 (Alexander et al., 2009), which employs a maximum likelihood-based clustering algorithm to estimate ancestry proportions of individuals from multilocus SNP data. The number of ancestral populations (K) varied from 1 to 25 to explore potential population stratification. The optimal K value was identified using the cross-validation (CV) error procedure implemented in ADMIXTURE, with the model exhibiting the lowest CV error considered as the best fit. Prior to analysis, SNPs were filtered for missing data and minor allele frequency (MAF), and linkage disequilibrium (LD) pruning was performed using PLINK

v1.9 (Purcell et al., 2007) to reduce marker redundancy. The resulting ancestry proportions (Q-matrix) were visualized as stacked bar plots using R v4.3.0.

Linkage disequilibrium (LD) decay was evaluated using PopLDdecay v3.41 (Zhang et al., 2019) following SNP pruning to remove highly correlated markers and ensure unbiased LD estimation. Pairwise LD (R^2) values were computed between SNPs across each chromosome, and the average R^2 was plotted against physical distance using R v4.3.0. Visualization was performed with the ggplot2 and ggsci packages using loess smoothing to describe the decay trend for each sub-genome (A, B, and D) as well as the whole genome (Wickham, 2016; Xiao, 2016).

Genome-Wide Association Analysis

Genome-wide association analyses were performed using four statistical models: the General Linear Model (GLM), Mixed Linear Model (MLM) (Yu et al., 2006), Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al., 2019), and the Fixed and Random Model Circulating Probability Unification (FarmCPU) (Liu et al., 2016). All analyses were conducted using the Genomic Association and Prediction Integrated Tool (GAPIT) (Lipka et al., 2012) implemented in the R environment (v4.4.2) within RStudio. GLM uses principal components as fixed effects, while the MLM, BLINK, and FarmCPU models are advanced multiple-locus models that iteratively identify associated markers as cofactors to improve statistical power and reduce Type I error rates. Model performance was evaluated through quantile–quantile (Q–Q) plots, which revealed that BLINK and FarmCPU outperformed GLM and MLM, consistent with previous studies (Yin et al., 2021). Both models demonstrated superior control of false positives and improved sensitivity in detecting true marker–trait associations (MTAs). Therefore, BLINK and FarmCPU were selected for the final association analysis to identify genomic regions associated with bacterial leaf streak (BLS) resistance. Visualization of GWAS results was conducted in RStudio using the CMplot package (Yin et al., 2021). Manhattan and Q–Q plots were generated based on the $-\log_{10}(P)$ transformation of p -values. The genome-wide significance thresholds were set at $-\log_{10}(P) \geq 4.0$ ($p = 0.05$) and $-\log_{10}(P) \geq 5.6$ ($p = 0.01$), following established criteria (Halder et al., 2023; Liu et al., 2017; Sidhu et al., 2019). Only MTAs surpassing the threshold in the combined environment (CEnv) were selected for further analyses.

Allelic Effects of MTAs on BLS Score and Marker Development for MAS

We dichotomized disease severity as “Resistant” for BLS scores 1–3 and “Susceptible” for BLS scores 7–9, excluding intermediate scores (4–6), and analyzed only lines with a non-missing value at each loci. For each allele of each MTA, we estimated carrier prevalence within each phenotypic group as the proportion of lines in that group with at least one copy of the allele (heterozygotes counted as carriers) and displayed paired bar plots with exact binomial 95% confidence intervals. Per SNP, the association between phenotype and genotype was tested with a 2×3 contingency analysis (Resistant/Susceptible × genotype): Pearson’s chi-square when all expected cell counts were ≥ 5 , otherwise the Fisher–Freeman–Halton exact test; two-sided P values are reported. To gauge the utility to conduct marker-assisted selection (MAS), we summarized per-allele effects as the difference in carrier prevalence between groups (Δ prevalence) and the carrier odds ratio for Resistant vs Susceptible lines with exact 95% confidence intervals from a 2×2 table. A SNP is considered promising for MAS when one allele shows clear enrichment in resistant lines (large positive Δ prevalence and the alternate allele shows the reciprocal pattern).

Results

Distribution of Bacterial Leaf Streak (BLS) Disease Severity across the RGON Wheat Population

A high uniformity was found between the experiments, as indicated by the analysis of variance (ANOVA), which revealed highly significant effects of genotype ($F = 77.62$, $P < 2 \times 10^{-16}$) and Replication ($F = 27.78$, $P = 1.26 \times 10^{-12}$), along with a low residual mean square (0.291), (Supplementary Table S1). The phenotypic distribution of Bacterial Leaf Streak (BLS) severity among the elite RGON wheat breeding lines displayed a continuous variation ranging from highly resistant (score = 1) to highly susceptible (score = 9) (Fig. 2A). Lines were categorized as highly resistant (scores 1–2), moderately resistant (scores 3–5), moderately susceptible (scores 6–7), and highly susceptible (scores 8–9). Out of 412 lines tested we report only 17 lines highly resistant to BLS (score 1–2), which accounts for about 4% of elite germplasm (Table 1). Overall, the majority of lines were in the moderately susceptible category (scores 6–7), representing approximately 35–40% of the population, indicating a predominance of intermediate susceptibility (Fig. 2B). Regional contributions to resistance and susceptibility varied substantially. Texas A&M (TX) and Colorado State University (CSU) breeding programs contributed the largest numbers of highly resistant lines (scores 1–2), comprising roughly 5% of the total population (20 lines per region), suggesting these programs harbor valuable alleles conferring strong BLS resistance. Kansas State University (KSU) and Texas A&M had the highest proportions of highly susceptible lines (scores 8–9), representing approximately 15% and 11% of the total lines, respectively, reflecting the coexistence of both resistant and susceptible germplasm in Texas, whereas KSU lines were predominantly susceptible. Other breeding programs, including North Dakota State University (NDSU), University of Nebraska Lincoln (NEB), Montana State University (MSU), South Dakota State University (SDSU), Oklahoma State University (OSU), Virginia Tech University (VA Tech), and BASF, contained relatively few highly susceptible lines and moderate numbers of moderately resistant to moderately susceptible lines (scores 3–7), indicating stable intermediate resistance across these populations. Notably, BASF lines were exclusively highly or moderately resistant, highlighting their potential as elite donors of durable BLS resistance alleles. OSU and VA lines, although limited in number, showed a mixture of moderately resistant and moderately susceptible responses, contributing to the overall phenotypic diversity. Overall, these results demonstrate substantial regional variation in BLS response among breeding programs. The CSU, NDSU, NEB, and BASF breeding programs emerged as key sources of resistance alleles. The broad spectrum of resistance and susceptibility observed within the Texas breeding program provides opportunities for both the immediate utilization of resistant lines and the improvement of susceptible materials. Collectively, this regional diversity offers a strong foundation for marker-assisted selection and genomic approaches aimed at developing hard winter wheat cultivars with durable resistance to BLS.

Table 1

Wheat breeding lines evaluated for bacterial leaf streak (BLS) response, including line designation, putative market class (HWW = hard white wheat; HRW = hard red wheat), source breeding program, and mean BLS disease rating. Disease ratings were based on a standardized scale, with lower values indicating greater resistance and higher values indicating increased susceptibility.

Line	Putative market class	Source (program)	BLS Disease Score (Mean)
22Nord-169	HWW	NDSU	1.83
23SeniorA-48	HWW	NDSU	2.00
23SeniorC-36	HWW	NDSU	1.33
VA21HRW-39	HRW	Virginia Tech	1.83
BASF-23-5		BASF	1.83
TX21M5111	HRW	TAMU-CS	1.00
TX21M5440	HRW	TAMU-CS	1.16
TX21M5477	HRW	TAMU-CS	1.66
TX21M5528	HRW	TAMU-CS	2.00
TX21M5629	HRW	TAMU-CS	1.66
CO20D036R	HRW	CSU	2.00
CO20D045R	HRW	CSU	1.83
CO20D068W	HWW	CSU	2.00
CO200034R	HRW	CSU	1.66
CO20SF227R	HRW	CSU	1.16
CO20SF163R	HRW	CSU	1.40
KS22U6864.1H.1R.R2.A1		USDA-ARS, KS	2.00

Genetic Diversity in Winter Wheat Genotypes and Population Structure Analysis

The 27,211 SNP markers collected from the 412 wheat breeding lines underwent a thorough filtering process (Minor Allele Frequencies (MAFs) of less than 0.05 or had more than 10% missing genotypic data), resulting in the selection of 15,368 high-quality SNP markers (supplementary file). Most of these SNP markers were distributed across the wheat A and B sub-genomes, accounting for 40% and 50% of the total, respectively, whereas the D-genome exhibited comparatively lower coverage at 10%. Chromosome 3B harbored the highest number of SNP markers, followed by 7A, 6B, and 2B (Fig. 3).

The genetic relationships among the genotypes were assessed using neighbor-joining phylogenetic clustering and principal component analysis (PCA) based on genome-wide SNP data. The phylogenetic tree (Fig. 4A) clearly divided the genotypes into five distinct clusters, represented by different colors, indicating genetic divergence within the collection. Each cluster contained genotypes that were more closely related to each other, while showing greater genetic distance from those in other clusters. The well-defined branching pattern and bootstrap support values further confirmed the reliability of the clustering. The PCA plot (Fig. 4B) further supported this population structure, with the first two principal components (PC1 and PC2) explaining 8.14% and 7.59% of the total genetic variation, respectively. Samples belonging to the same phylogenetic cluster were positioned closely in the PCA space, confirming the consistency between both analyses. The smooth color gradient observed across clusters suggested partial admixture or gradual genetic

differentiation among subpopulations rather than complete isolation. Together, these analyses demonstrate that the studied population is genetically structured into multiple clusters, reflecting underlying genetic diversity and likely explained by geographic origin of each represented state and its related breeding program.

Population genetic structure among the genotypes was inferred using the ADMIXTURE algorithm based on genome-wide SNP data. The cross-validation (CV) error values gradually decreased with increasing K values and reached the minimum at K = 5, after which the error began to rise (Fig. 5A). This showed that five ancestral genetic clusters (K = 5) best described the population structure within the dataset. The ancestry proportion plot (Fig. 5B) illustrates the membership coefficient of each genotype to the inferred ancestral populations, represented by distinct colors. Each vertical bar corresponds to an individual genotype, and the color proportions within each bar indicate the estimated fraction of ancestry derived from each of the five clusters. The results revealed a high degree of genetic admixture, suggesting that the genotypes share founder breeding lines and likely common parental lines in the pedigree of evaluated lines. However, some degree of population sub-structuring was evident, isolated from certain regions such as Virginia Tech and NDSU, which showed predominant affiliation to specific ancestral clusters, indicating regional wheat production class differences as Virginia Tech predominantly grows soft white wheat and NDSU grows hard red winter wheat and/or regional genetic differentiation. In contrast, genotypes from breeding programs of KSU, SDSU, and Texas A&M exhibited mixed ancestry, reflecting sharing of genotypes, historical gene flow or recent recombination events among subpopulations and predominantly same class of wheat is grown i.e. hard red winter wheat. Overall, ADMIXTURE analysis delineated the population into five genetically distinct but interrelated subpopulations, consistent with the results from phylogenetic tree and PCA analyses. The observed admixture patterns suggest a complex evolutionary relationship among genotypes, possibly shaped by geographical distribution, host adaptation, and historical genetic exchange. The kinship analysis further segregated the panel into three distinct clusters, indicating a notable genetic disparity among the included accessions. The phylogenetic tree supports this division, while the left side of the heatmap confirms the structural analysis results. Additionally, the varying intensities of color in the heatmap highlight regions of heightened linkage disequilibrium (**Supplementary Fig. 1**).

Linkage Disequilibrium Decay

Genome-wide LD decay analysis revealed clear differences in the extent and rate of LD decay among the three sub-genomes of hexaploid wheat (Fig. 6). The A sub-genome showed the highest initial LD ($r^2 \approx 0.12$) and the slowest decay rate, maintaining substantial LD across larger genomic distances. LD in the A sub-genome decayed to half of its initial value at approximately 4–5 Mb, suggesting limited historical recombination and the presence of longer haplotype blocks. This slower decay may be attributed to stronger selection pressure, lower effective recombination, or reduced nucleotide diversity within the A sub-genome. The B sub-genome showed an intermediate LD decay pattern, with an initial $r^2 \approx 0.09$ that gradually declined to near-background levels (~ 0.03) at about 6–7 Mb. This moderate decay rate indicates a balance between recombination and selection, implying that the B-genome has experienced moderate genetic reshuffling during evolution. In contrast, the D sub-genome showed apparently rapid LD decay (~ 0.03 to ~ 0.01 within 1–2 Mb), this is likely an artifact because of its low SNP density and limited genetic diversity. The small number of markers and reduced haplotype variation can give the appearance of faster decay, despite overall lower recombination and diversity in the D sub-genome (Pang et al., 2020; Chen et al., 2019). When all sub-genomes were combined, the whole-genome LD decay curve closely resembled that of the B sub-genome, with LD decaying to background levels around 5 Mb. The LD decay cutoff, defined as the distance at which r^2 falls below 0.1 or reaches the background level (~ 0.02 – 0.03), therefore indicates that linkage disequilibrium across the wheat genome generally breaks down between 4 and 6 Mb.

Association Mapping for BLS Response

The genome-wide association study (GWAS) was conducted using the BLINK and FarmCPU models to identify significant marker–trait associations (MTAs) linked to bacterial leaf streak (BLS) disease resistance in wheat. Both models demonstrated strong statistical performance, as evidenced by their well-distributed quantile–quantile (Q–Q) plots (Fig. 7),

indicating effective control of false positives and an ability to capture true genetic signals. Compared to traditional models such as GLM and MLM, the BLINK and FarmCPU models were more powerful in detecting associations in a population with complex structure, as they integrate both fixed and random effects while minimizing the confounding influences of population stratification. The BLINK and FarmCPU models jointly identified seven significant genomic regions associated with BLS resistance, located on chromosomes 1B, 2B, 3B, 3D, 4A, 4B, and 6B (Table 2).

Table 2
Summary of significant SNPs associated with QTLs for resistance to Bacterial Leaf Streak (BLS), identified from the Fixed and Random Model Circulating Probability Unification (FarmCPU) and Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) GWAS model.

SNPs	Chr	Pos (bp)	Major Allele	Minor Allele	P value	MAF	PVE (%)	Effect	GWAS model
S2B_175365127	2B	175365127	C	A	1.44E-06	0.14	16.56	0.51	BLINK+ FarmCPU
S4B_322341647	4B	322341647	A	T	1.78E-06	0.05	15.31	-0.80	BLINK
S1B_372870844	1B	372870844	G	T	1.36E-06	0.17	13.8	1.02	BLINK
S6B_548788571	6B	548788571	G	A	1.97E-06	0.07	10.39	-0.74	BLINK+ FarmCPU + GLM
S3B_625137880	3B	625137880	T	C	1.46E-07	0.24	1.13	0.45	FarmCPU + GLM
S4A_440017300	4A	440017300	T	A	1.51E-06	0.05	1.11	0.78	FarmCPU
S3D_618866799	3D	618866799	G	C	3.91E-06	0.07	0.04	0.68	BLINK+ FarmCPU + GLM

Among these, three MTAs were common to both models, highlighting their consistency and robustness in detecting true associations. The phenotypic variance explained (PVE) by these markers ranged from 0.04% to 16.56%, indicating that both models were capable of capturing loci with small to larger effects. Notably, the BLINK model detected SNPs with higher PVE values, such as S2B_175365127 (16.56%) and S4B_322341647 (15.31%), suggesting its superior sensitivity in identifying loci with stronger genetic effects. In contrast, the FarmCPU model was more efficient in detecting loci with smaller effects, such as S3B_625137880 (1.13%) and S4A_440017300 (1.11%), demonstrating its strength in uncovering polygenic contributions to resistance. Together, these findings indicate that the BLINK model tends to identify major-effect loci with strong associations, while the FarmCPU model effectively captures both major- and minor-effect loci, leading to a more comprehensive understanding of the genetic basis of BLS resistance. The complementary strengths of these two models ensure a balanced detection of key genomic regions influencing the trait. Overall, the variation explained by these significant SNPs supports the polygenic nature of BLS resistance and underscores the potential of BLINK and FarmCPU as reliable tools for dissecting complex traits and advancing marker-assisted breeding in wheat. Overall, the seven loci represented the unique genomic regions associated with BLS resistance.

Allele Analysis for MTAs identified

To validate the MTAs for marker development, the allelic effects of the MTAs on BLS response were estimated using contingency tables. Briefly, the BLS score ranging 1–2 was factorized as “highly resistant (HR)”, score 3–5 as “moderately resistant (MR)”, score 6–7 as “moderately susceptible (MS)”, which is short for moderate resistance or susceptibility, score 8–9 as “highly susceptible (HS)”. To simplify the visualization of the allelic effect of each MTA, “HR” and “MR”

categories were merged and classified as “resistant lines”. Likewise, “HS” and “MS” were merged and classified as susceptible lines. For the seven MTAs detected, we validated the effect of alleles on the BLS disease phenotype by Fisher–Freeman–Halton exact test. Out of the seven MTA’s detected, three MTAs, *S1B_372870844* (T allele), *S2B_175365127* (A allele), and *S4B_322341647* (A allele) allelic effect of resistance to BLS disease were significant at $p < 0.01$ (Fig. 8). A fourth MTA, *6B_548788571* (G allele), allelic effect was significant at $p < 0.05$. The *S4B_322341647* (A allele) and *6B_548788571* (G allele), were present in all the seventeen highly resistant lines. The T allele of *S4B_322341647* was found almost exclusively in susceptible lines, making it useful for identifying susceptibility. Thus, there is statistical confidence in the odds of the allele increasing resistance to BLS disease and these four MTAs can be used with confidence in the development of KASP markers for pre-breeding and Marker-assisted selection.

Additive Effect of identified Marker Trait Association Loci and potential of Pyramiding

The seven MTA’s were then evaluated for additive effect of MTA’s with resistance alleles and MTA’s with alternate alleles on the BLS disease severity of wheat (Fig. 9). The mean disease rating score was inversely correlated with the addition of number of resistant loci with favorable alleles, indicating that with more favorable alleles the disease score reduced ($p = 4.6 \times 10^{-20}$). In this study, we observed that with addition of each resistance allele the disease severity dropped by 1.449 units. The best combination of MTAs that were able to provide the least disease score of 1, were observed in wheat line TX21M5111 and TX21M5440, where both carried the resistant alleles at the MTAs, *S1B_372870844*, *S4B_322341647* and *6B_548788571* (supplementary Table 3). These three MTAs provided a total PVE of $\sim 40\%$ in this study. Similarly, the additive effect of susceptible alleles showed a positive correlation, showing that with every addition of susceptible allele, the BLS disease score increased by 1.445 units towards susceptibility at $p = 3.6 \times 10^{-20}$ (Supplementary Fig. 2).

Linkage Disequilibrium blocks determination at the significant MTAs

The identified loci in this study explained 0.04 – 16.56% of phenotypic variance, representing major and minor MTAs for BLS disease resistance in elite bread wheat. Although no universal PVE cutoff exists, for our study, we focused on the loci exceeding 10% that are typically considered biologically significant and highlight key contributors to trait variation (Manolio et al., 2009; Visscher et al., 2017). Reporting PVE alongside statistical significance provides a clear measure of each locus's contribution to the genetic architecture (Lee et al., 2011; Yang et al., 2010). A high LD (> 0.8) was found within a 10kb block for MTAs *S1B_372870844*, *S3B_625137880* and *S4B_322341647* (Fig. 10).

DISCUSSION

Bacterial leaf streak (BLS) is a significant concern for various major crops, including wheat, barley, maize, Triticale, rice, and Brassica ssp. (El Attari et al., 1996; Qiu et al., 2020; Ritzinger et al., 2023; Tang et al., 2000; Wen et al., 2018). Wheat yields have been reported to be reduced significantly due to the BLS incidence and disease severity in the last decade, particularly in the northern great plains of the US (Friskop et al., 2023). The development of disease-resistant cultivars has emerged as a key strategy for managing BLS, particularly in the absence of effective chemical controls (Milus & Mirlohi, 1995). However, progress in breeding to develop resistant cultivars has been impeded by the complex inheritance patterns of BLS resistance, the large genome size of wheat, limited understanding and or studies of resistance mechanisms, and the absence of suitable molecular markers to use for the development of BLS resistant varieties. However, recent advances made in identifying several major genes and quantitative trait loci (QTLs) associated with BLS resistance in wheat (Ramakrishnan et al., 2019) and Triticale (Wen et al., 2018) have begun to address the key knowledge gaps. Previous research by Duveiller et al. (1992) conducted at CYMMYT involved extensive crossbreeding among five wheat parent varieties, revealing mapping of five genes (*Bls1*, *Bls2*, *Bls3*, *Bls4*, and *Bls5*) in BLS resistance. Studies conducted by Tillman et al. (1996), Kandel et al. (2012), and Adhikari et al. (2012) have demonstrated partial BLS resistance in wheat, with most genetic investigations focusing on spring wheat varieties and only a few addressing winter wheat (Ramakrishnan et al., 2019; Tillman et al., 1996). In this study, we evaluated 412 elite hard winter wheat elite breeding lines

from the contemporary US, adapted to diverse U.S. Agro-climatic regions, for their response to bacterial leaf streak under greenhouse-controlled conditions. The use of contemporary wheat accessions from a larger geographical range is valuable, as it encompasses the warmer and more humid southern states that may harbor bacterial resistance. The observed phenotypic distributions exhibited a range from resistant to susceptible genotypes and the majority leaned towards moderate susceptibility, consistent with the previous observations of the majority in susceptibility of wheat germplasm to BLS (Adhikari et al., 2012; Duveiller et al., 1992; Tillman et al., 1996), which also aligns with findings in rice-*Xanthomonas oryzae* pv *oryzae* pathosystem (Poulin et al., 2014; Tang et al., 2000).

The evaluation of BLS severity among the RGON wheat accessions revealed a wide phenotypic range, from highly resistant (score = 1) to highly susceptible (score = 9) (Fig. 3A). Most accessions were moderately susceptible (score = 6), indicating a predominance of partial susceptibility within the population (Fig. 2B). We identified 17 accessions with strong resistance to BLS, which accounted for 4% of the genotypes tested in this study. Among the 17 resistant accessions, the most resistant line from Texas A&M was TX21M5111, with an average BLS disease score of 1 (highly resistant). Highest resistant from Texas supports our hypothesis that the agro climatic zone of higher humidity and warm temperatures, could have led to selection and breeding of disease resistant accessions. Among the 412 accessions, the Texas A&M (TX) and Colorado State breeding programs (CSU) contributed the greatest number of resistant lines (scores 1–3), also suggesting the presence of valuable resistance alleles within their breeding programs that could be utilized in breeding for durable BLS resistance. In contrast, Kansas (KSU) had the highest numbers of highly susceptible lines (scores 7–9), with Kansas showing a predominance of susceptibility, highlighting the need for introgression of resistance alleles from other regions. Accessions from NDSU, UNL, and MSU exhibited moderate to intermediate resistance, indicating stable resistance sources that may perform consistently across environments and efforts for host resistance deployment. The broad diversity within Texas germplasm, encompassing both resistant and susceptible lines, makes it an important source for both resistance enhancement and breeding advancement.

Although, we tested the four statistical models for genome-wide association mapping, we identified seven quantitative trait loci (QTLs) associated with resistance to BLS in winter wheat by using two complementary models BLINK and FarmCPU, as evidenced by their well-distributed quantile–quantile (Q–Q) plots, indicating effective control of false positives and an ability to capture true genetic signals. These loci were distributed across seven chromosomal regions: 2B, 4B, 1B, 6B, 3B, 4A and 3D, explaining 16.56%, 15.31%, 13.8%, 10.39%, 1.13%, 1.11%, and 0.044% of the total Phenotypic Variation Explained (PVE), respectively. The 2B, 6B, 3B and 3D QTLs were detected by both BLINK and FarmCPU models and therefore are stable and consistent with BLS resistance in different statistical models. The MTA's effect ranged from 1.02 to -0.80, indicating most markers detected are major loci for the BLS disease resistance, unlike the previous studies (Tillman et al. (1996), Kandel et al. (2012), and Adhikari et al. (2012), (Thapa et al., 2025)). Among the most resistant 17 lines detected in our study (Table 2), they carried 5 to 2 resistant alleles of the MTAs detected in this study. All the 17 highly resistant lines reported in this study, carried the *S4B_322341647* (~ 15% PVE) and *S6B_548788571* (~ 10% PVE). Four lines, CO200034R, CO20SF227R, TX21M5477 and CO20D045R contained the other two major QTLs at *S2B_175365127* (~ 17% PVE) and *S1B_372870844* (~ 14% PVE). In addition to these four MTAs, one highly resistant line, CO20SF163R carries a fifth resistant MTAs, *S3B_625137880* (~ 1%PVE), indicating a presence of good resistant materials with major QTLs for deployment into adapted cultivars for BLS resistance.

Four MTAs detected in this study were found to be highly significant for their allelic effect and therefore are very valuable for application in marker-assisted selection for BLS resistance. The additive effect of the stacking of favorable alleles indicates that pyramiding of the MTAs would lead to the deployment of durable BLS resistance in wheat. Based on the LD blocks estimation and comparison with previous studies, we estimated with varying degrees of confidence whether the identified SNPs that are significantly associated with BLS resistance overlap with previously reported QTLs or are novel (Table 3). Among the seven loci detected, the QTLs 4A overlapped with regions previously reported in independent GWAS analyses for BLS resistance in winter wheat (Ramakrishnan et al., 2019) and spring wheat (Adhikari et al., 2012). Additionally, earlier studies by Adhikari et al. (2012) and Kandel et al. (2015b) have reported genomic associations for BLS

resistance on chromosomes 4B, 6B, and 2A, which are consistent with loci identified in our results. Importantly, we discovered a novel QTL on chromosome 2B, which has not been reported in prior studies, indicating a potentially unique resistance source in the SDSU germplasm.

Table 3
Comparison of significant SNPs identified in this study with previously published QTL for bacterial leaf streak (BLS) resistance in wheat.

SNP	Chr	Position (bp)	Published BLS QTL / Marker(s)	Reference Position (Mb)	Difference (Mb)	Reference(s)	Likely Relation
S2B_175,365,127	2B	175,365,127	No major or consistently validated BLS QTL reported on distal 2BL	17.29	158.07	(Gurung et al., 2014)	Novel locus
S4B_322,341,647	4B	322,341,647	Sporadic associations reported on 4B, but lacking consistency and positional agreement			This study	Uncertain; likely novel
S1B_372,870,844	1B	372,870,844	BLS QTL primarily reported on 1A or proximal 1BS; no validated QTL at this position	9.54	363.33	(Ramakrishnan et al., 2019)	Novel
S6B_548,788,571	6B	548,788,571	BLS-associated regions previously reported on 6B; however, reported QTLs are in different chromosomal intervals	19.26	529.53	(Gurung et al., 2014; Kandel et al., 2015b)	Novel
S3B_625,137,880	3B	625,137,880	Major BLS resistance QTL reported on chromosome 3B, located in the upper/proximal region	50.00	575.14	(Acharya et al., 2024)	Novel
S4A_440,017,300	4A	440,017,300	Minor BLS-associated regions reported on 4A, but not co-localized with this distal position	200.00 / 144.37	240.02 / 295.65	(Acharya et al., 2024; Ramakrishnan et al., 2019)	Possible novel locus
S3D_618,866,799	3D	618,866,799	None			This study	Novel

The MTAs identified in this study show varying relationships to previously reported BLS resistance MTAs. The marker *S3B_625,137,880* on chromosome 3B is located in the distal region of 3B; although a major BLS QTL has been reported on

3B (Acharya et al., 2024), it resides 4Mb in the upper/proximal region, indicating that this SNP likely represents a distinct locus. Similarly, *S6B_548,788,571* on chromosome 6B lies on a chromosome previously associated with BLS resistance, including GWAS signals reported by (Gurung et al., 2014) and the SSR marker *Xbarc134* reported by (Kandel et al., 2015b); however, the lack of positional concordance suggests this SNP may represent an independent resistance signal with medium confidence. In contrast, *S2B_175,365,127* (chromosome 2B), *S4A_440,017,300* (chromosome 4A), *S4B_322,341,647* (chromosome 4B), and *S1B_372,870,844* (chromosome 1B) appear to be novel loci, as they have not been consistently linked to BLS resistance in prior studies, though minor associations with related leaf-disease traits have been reported (Acharya et al., 2024; Adhikari et al., 2012). The SNP *S3D_618,866,799* on chromosome 3D also represents a novel candidate, as no stable BLS QTLs have been documented on this chromosome. Overall, these findings highlight both previously identified chromosomes harboring BLS resistance and several putative novel loci that contribute to resistance in hard winter wheat.

The primary aim of conducting genome-wide association mapping of BLS resistance was to identify the specific QTLs and underlying genes responsible for conferring resistance to BLS. Our findings will advance understanding of the complex pathosystem. The SNPs found to be associated with BLS resistance will streamline the process of selecting resistant individuals through backcrossing or utilizing double haploid populations, thus incorporating favorable alleles into deployment of high-yielding and disease-resistant winter wheat cultivars.

CONCLUSIONS

This study identified seven genomic regions associated with the resistance against bacterial leaf streak disease within the contemporary elite winter wheat accessions tested. The MTA's identified were responsible for upto ~16% phenotypic variance. This study identified 17 highly resistant breeding lines and determined that they carry up to two and as many as five favorable alleles conditioning resistance at the detected regions. Furthermore, four MTAs had a significant allelic effect on the BLS disease score, and those MTA's favorable alleles were additive, indicating their applicability for pyramiding in breeding schemes, the development of KASP assays for marker-assisted selection and the potential for deploying durable BLS resistance in wheat. Moreover, the seventeen BLS disease resistant genotypes identified herein offer a source of elite breeding materials that are already adapted to the Northern Great Plains region of the United States, which and can lead to the accelerated deployment of BLS resistance in regionally adapted, high yielding varieties.

Declarations

Funding

This study was carried out with the support of the Agricultural Experimentation Funding SD00H754, SD00H717, South Dakota Wheat Commission funding SA2400069 and SA2500056 and ARS Wheat CRIS Project #3042-21000-035-000D.

Conflict of interests

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

MA: Conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and tables, authored first drafts of the paper, **JB:** coordinator of RGON **HA:** computational analysis/ Figures, **PS, AB, KJ,** and **GB:** generated the genotyping data. **MA:** provided seed of the 2023 RGON genotypes and genotypic data. **KG:** Conceived and designed the experiments, **SS:** Conceived and designed the experiments/ computational analysis/ authored first drafts of the paper, **GA:** Conceived and designed the experiments,

analyzed the data, contributed reagents/materials/analysis tools, prepared figures, and tables, authored first drafts of the paper/ conceptualized the idea. All the authors read, reviewed, edited, and approved the manuscript.

Data Availability

The phenotyping and genotyping data are available as supplementary information.

References

1. Friskop A, Green A, Ransom J, Liu Z, Knodel J, Hansen B, Halvorson J, Lux L: Increase of bacterial leaf streak in hard red spring wheat in North Dakota and yield loss considerations. *Phytopathology*® 2023, 113(11):2103-2109.
2. Pandey P, Irulappan V, Bagavathiannan MV, Senthil-Kumar M: Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Frontiers in plant science* 2017, 8:537.
3. Ritzinger MG, Smith KP, Case AJ, Wodarek JR, Dill-Macky R, Curland RD, Steffenson BJ: Sources of bacterial leaf streak resistance identified in a diverse collection of barley germplasm. *Plant disease* 2023, 107(3):802-808.
4. Sapkota S, Mergoum M, Liu Z: The translucens group of *Xanthomonas translucens*: Complicated and important pathogens causing bacterial leaf streak on cereals. *Molecular plant pathology* 2020, 21(3):291-302.
5. Heiden N, Roman-Reyna V, Curland R, Dill-Macky R, Jacobs J: Barley-infecting *Xanthomonas translucens* genomic groups are globally distributed. In: *PHYTOPATHOLOGY: 2022: AMER PHYTOPATHOLOGICAL SOC 3340 PILOT KNOB ROAD, ST PAUL, MN 55121 USA; 2022: 4-4*.
6. Ramakrishnan SM, Sidhu JS, Ali S, Kaur N, Wu J, Sehgal SK: Molecular characterization of bacterial leaf streak resistance in hard winter wheat. *PeerJ* 2019, 7:e7276.
7. Kandel YR, Glover KD, Tande CA, Osborne LE: Evaluation of spring wheat germplasm for resistance to bacterial leaf streak caused by *Xanthomonas campestris* pv. *translucens*. *Plant Disease* 2012, 96(12):1743-1748.
8. Hong E, Bankole IA, Zhao B, Shi G, Buck JW, Feng J, Curland RD, Baldwin T, Chapara V, Liu Z: DNA markers, pathogenicity test, and multilocus sequence analysis to differentiate and characterize cereal-specific *Xanthomonas translucens* strains. *Phytopathology*® 2023, 113(11):2062-2072.
9. Silva IT, Rodrigues FÁ, Oliveira JR, Pereira SC, Andrade CCL, Silveira PR, Conceição MM: Wheat resistance to bacterial leaf streak mediated by silicon. *Journal of Phytopathology* 2010, 158(4):253-262.
10. Duveiller E, Bragard C, Maraite H, eds: Bacterial leaf streak and black chaff caused by *Xanthomonas translucens*. In.: International Maize and Wheat Improvement Center, Estado de México, Mexico; 1997: 25-47.
11. Shohanipor T, Sharifi R, Hosseini S: Efficacy Assessment of Some Copper Compounds on the Bacterial Leaf Streak Pathogen, *Xanthomonas translucens*. *European Journal of Plant Pathology* 2025:1-16.
12. Tillman B, Harrison S, Russin J, Clark C: Relationship between bacterial streak and black chaff symptoms in winter wheat. *Crop science* 1996, 36(1):74-78.
13. Duveiller E, van Ginkel M, Thijssen M: Genetic analysis of resistance to bacterial leaf streak caused by *Xanthomonas campestris* pv. *undulosa* in bread wheat. *Euphytica* 1992, 66:35-43.
14. 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, 5(1).
15. Wen A, Jayawardana M, Fiedler J, Sapkota S, Shi G, Peng Z, Liu S, White FF, Bogdanove AJ, Li X: Genetic mapping of a major gene in triticale conferring resistance to bacterial leaf streak. *Theoretical and applied genetics* 2018, 131:649-658.
16. Kandel YR, Glover KD, Osborne LE, Gonzalez-Hernandez JL: Mapping quantitative resistance loci for bacterial leaf streak disease in hard red spring wheat using an identity by descent mapping approach. *Euphytica* 2015, 201:53-65.

17. Tang D, Wu W, Li W, Lu H, Worland A: Mapping of QTLs conferring resistance to bacterial leaf streak in rice. *Theoretical and applied genetics* 2000, 101:286-291.
18. HE W-a, HUANG D-h, LI R-b, QIU Y-f, SONG J-d, YANG H-n, ZHENG J-x, HUANG Y-y, LI X-q, Chi L: Identification of a Resistance Gene bls1 to Bacterial Leaf Streak in Wild Rice *Oryza rufipogon* Griff. *Journal of Integrative Agriculture* 2012, 11(6):962-969.
19. Zhu C, Gore M, Buckler ES, Yu J: Status and prospects of association mapping in plants. *The plant genome* 2008, 1(1).
20. Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J: Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nature communications* 2011, 2(1):467.
21. Begum H, Spindel JE, Lalusin A, Borromeo T, Gregorio G, Hernandez J, Virk P, Collard B, McCouch SR: Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (*Oryza sativa*). *PloS one* 2015, 10(3):e0119873.
22. Bordes J, Ravel C, Le Gouis J, Lapierre A, Charmet G, Balfourier F: Use of a global wheat core collection for association analysis of flour and dough quality traits. *Journal of Cereal Science* 2011, 54(1):137-147.
23. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB: A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics* 2006, 38(2):203-208.
24. Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM: Mixed linear model approach adapted for genome-wide association studies. *Nature genetics* 2010, 42(4):355-360.
25. Pritchard JK, Rosenberg NA: Use of unlinked genetic markers to detect population stratification in association studies. *The American Journal of Human Genetics* 1999, 65(1):220-228.
26. Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000, 155(2):945-959.
27. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* 2006, 38(8):904-909.
28. Letta T, Olivera P, Maccaferri M, Jin Y, Ammar K, Badebo A, Salvi S, Noli E, Crossa J, Tuberosa R: Association mapping reveals novel stem rust resistance loci in durum wheat at the seedling stage. *The Plant Genome* 2014, 7(1):plantgenome2013.2008.0026.
29. Muleta KT, Rouse MN, Ryneerson S, Chen X, Buta BG, Pumphrey MO: Characterization of molecular diversity and genome-wide mapping of loci associated with resistance to stripe rust and stem rust in Ethiopian bread wheat accessions. *BMC plant biology* 2017, 17:1-20.
30. Raboin L-M, Ballini E, Tharreau D, Ramanantsoanirina A, Frouin J, Courtois B, Ahmadi N: Association mapping of resistance to rice blast in upland field conditions. *Rice* 2016, 9:1-12.
31. Xiao Y, Liu H, Wu L, Warburton M, Yan J: Genome-wide association studies in maize: praise and stargaze. *Molecular plant* 2017, 10(3):359-374.
32. Ahmad M, Alhusays A, Galvin C, Torres M, Fitzpatrick E, Glover K, Solanki S, Ameen G: Bacterial leaf streak disease resistance loci identification in winter wheat. In: *PHYTOPATHOLOGY: 2024: AMER PHYTOPATHOLOGICAL SOC 3340 PILOT KNOB ROAD, ST PAUL, MN 55121 USA; 2024: 105-105*.
33. Ahmad M, Pawar T, Solanki S, Glover K, Ameen G: An Effective Method for Bacterial Leaf Streak Disease Severity Estimation in Controlled and Field Environments in Small Grains. 2025.
34. Gomez KA, Gomez AA: Statistical procedures for agricultural research: John wiley & sons; 1984.
35. Bernardo A, St. Amand P, Le HQ, Su Z, Bai G: Multiplex restriction amplicon sequencing: a novel next-generation sequencing-based marker platform for high-throughput genotyping. *Plant biotechnology journal* 2020, 18(1):254-265.

36. 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, 23(19):2633-2635.
37. Browning SR, Browning BL: Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics* 2007, 81(5):1084-1097.
38. Perrier X: DARwin software. <http://darwin.cirad.fr/darwin> 2006.
39. Alexander DH, Novembre J, Lange K: Fast model-based estimation of ancestry in unrelated individuals. *Genome research* 2009, 19(9):1655-1664.
40. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ: PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics* 2007, 81(3):559-575.
41. Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L: PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 2019, 35(10):1786-1788.
42. Wickham H: Data analysis. In: *ggplot2: elegant graphics for data analysis*. Springer; 2016: 189-201.
43. Xiao N: ggsci: scientific journal and Sci-Fi themed color palettes for 'ggplot2'. (No Title) 2016.
44. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z: BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 2019, 8(2):giy154.
45. Liu X, Huang M, Fan B, Buckler ES, Zhang Z: Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS genetics* 2016, 12(2):e1005767.
46. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z: GAPIT: genome association and prediction integrated tool. *Bioinformatics* 2012, 28(18):2397-2399.
47. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, Yuan X, Zhu M, Zhao S, Li X: rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, proteomics & bioinformatics* 2021, 19(4):619-628.
48. Sidhu JS, Ramakrishnan SM, Ali S, Bernardo A, Bai G, Abdullah S, Ayana G, Sehgal SK: Assessing the genetic diversity and characterizing genomic regions conferring Tan Spot resistance in cultivated rye. *PLoS One* 2019, 14(3):e0214519.
49. Liu H, Luo X, Niu L, Xiao Y, Chen L, Liu J, Wang X, Jin M, Li W, Zhang Q: Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. *Molecular Plant* 2017, 10(3):414-426.
50. Halder J, Gill HS, Zhang J, Altameemi R, Olson E, Turnipseed B, Sehgal SK: Genome-wide association analysis of spike and kernel traits in the US hard winter wheat. *The Plant Genome* 2023, 16(1):e20300.
51. Pang Y, Liu C, Wang D, Amand PS, Bernardo A, Li W, He F, Li L, Wang L, Yuan X: High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Molecular Plant* 2020, 13(9):1311-1327.
52. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A: Finding the missing heritability of complex diseases. *Nature* 2009, 461(7265):747-753.
53. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J: 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics* 2017, 101(1):5-22.
54. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW: Common SNPs explain a large proportion of the heritability for human height. *Nature genetics* 2010, 42(7):565-569.
55. Lee SH, Wray NR, Goddard ME, Visscher PM: Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics* 2011, 88(3):294-305.

56. El Attari H, Sarrafi A, Garrigues S, Dechamp-Guillaume G, Barrault G: Diallel analysis of partial resistance to an Iranian strain of bacterial leaf streak (*Xanthomonas campestris* pv. *cerealis*) in wheat. *Plant pathology* 1996, 45(6):1134-1138.
57. Qiu Y, Kaiser C, Schmidt C, Broders K, Robertson AE, Jamann TM: Identification of quantitative trait loci associated with maize resistance to bacterial leaf streak. *Crop Science* 2020, 60(1):226-237.
58. Milus E, Mirlohi A: Survival of *Xanthomonas campestris* pv. *translucens* between successive wheat crops in Arkansas. *Plant disease* 1995, 79(3):263-265.
59. Poulin L, Raveloson H, Sester M, Raboin L-M, Silué D, Koebnik R, Szurek B: Confirmation of bacterial leaf streak caused by *Xanthomonas oryzae* pv. *oryzicola* on rice in Madagascar. *Plant Disease* 2014, 98(10):1423-1423.
60. Thapa S, Singh M, Gill H, Ali S, Halder J, Koupal D, Sidhu JS, Sehgal S: Identification of novel sources and genetic mapping for Bacterial Leaf Streak resistance in a geographically diverse panel of wheat. *Plant Disease* 2025(ja).
61. Kandel YR, Glover KD, Osborne LE, Gonzalez-Hernandez JL: Mapping quantitative resistance loci for bacterial leaf streak disease in hard red spring wheat using an identity by descent mapping approach. *Euphytica* 2015, 201(1):53-65.
62. 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, 9(9):e108179.
63. Acharya K, Liu Z, Schachterle J, Kumari P, Manan F, Xu SS, Green AJ, Faris JD: Genetic mapping of QTLs for resistance to bacterial leaf streak in hexaploid wheat. *Theoretical and Applied Genetics* 2024, 137(12):265.

Figures

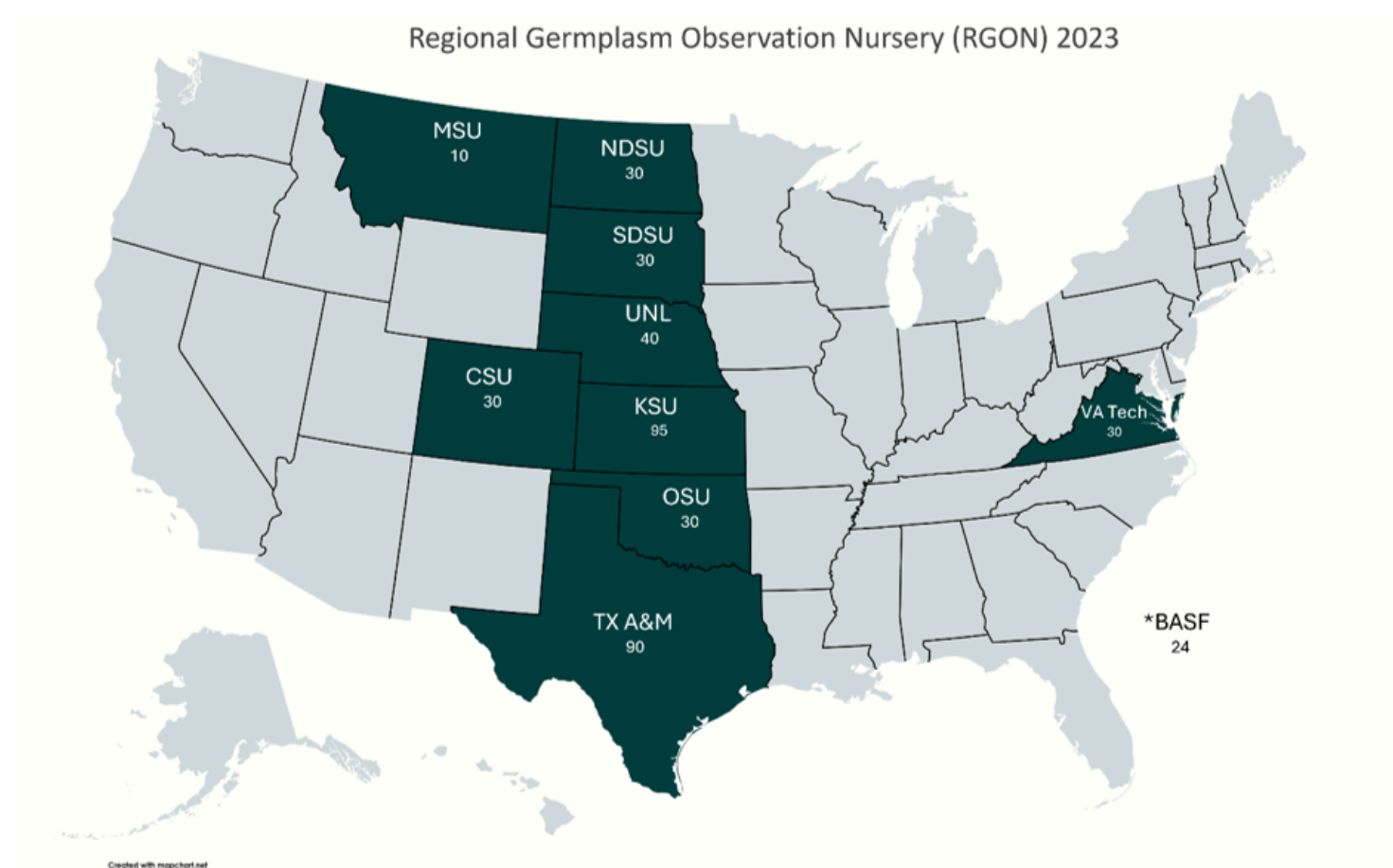


Figure 1

The distribution of the 412 elite winter wheat lines evaluated in this study for bacterial leaf streak disease resistance from the different wheat-producing states of the US. Highlighted dark green color with the respective numbers represents the experimental lines developed in the breeding programs within each state in the US.

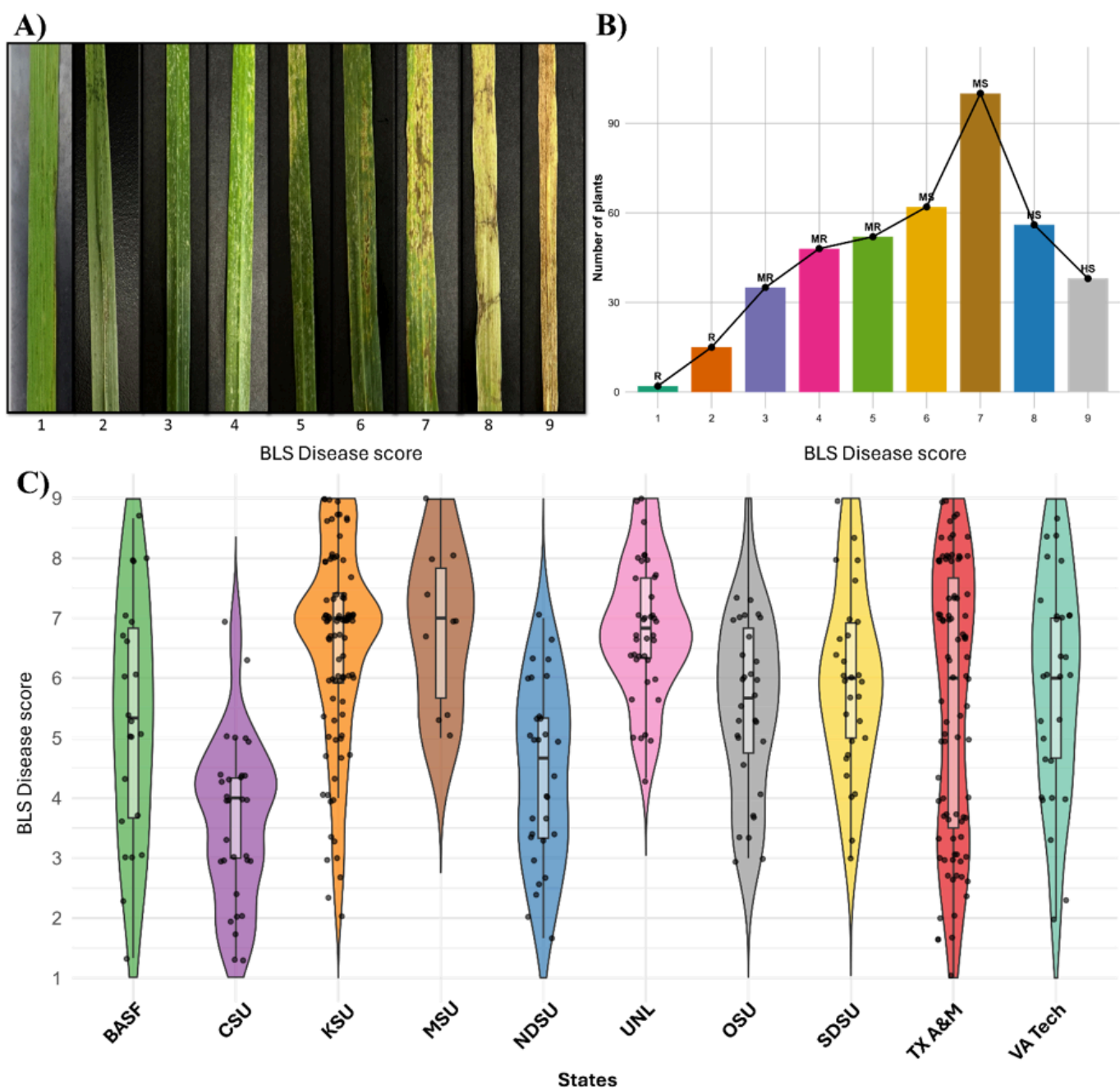


Figure 2

Bacterial leaf streak disease severity rating distribution of the 412 elite hard winter wheat genotypes. **(A)** Representative BLS disease rating scale on wheat leaves showing a continuous range of BLS symptom severity from resistant (R) to highly susceptible where 1 indicates highly resistant and 9 indicates highly susceptible **(B)** Bar plot with overlaid line graphs illustrating the frequency distribution of wheat lines tested, across BLS rating scale of the nine continuous categories (R-HS), **R**= resistant, **MR**= Moderately resistant, **MS**= Moderately susceptible, and **HS**=Highly susceptible. **(C)**

Violin plots with embedded boxplots and individual data points, showing the distribution of BLS disease severity values across regions (BASF, CO, KS, MT, ND, NEB (UNL), OK, SD, TX (TXA&M), VA).

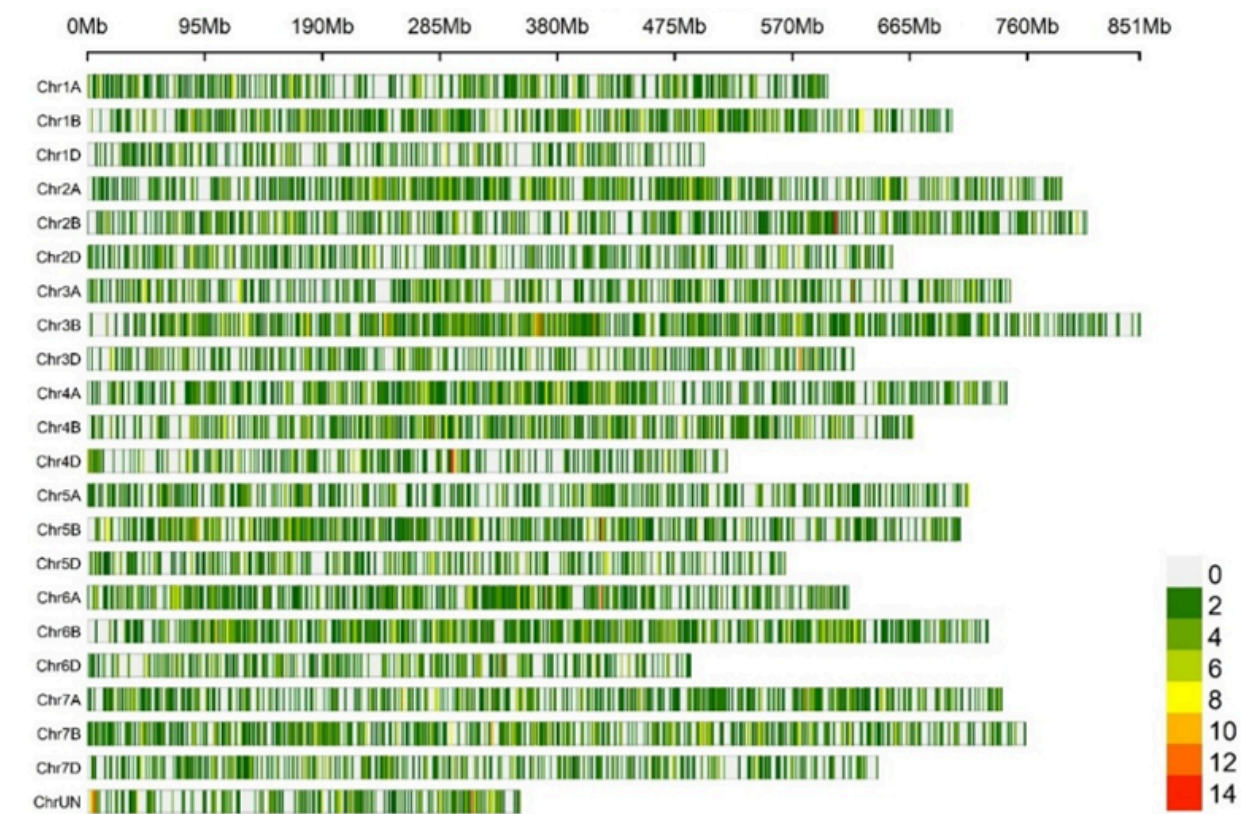


Figure 3

Genetic Marker distribution on the 21 chromosomes (A, B, and D) of the elite winter wheat genotypes used in this study.

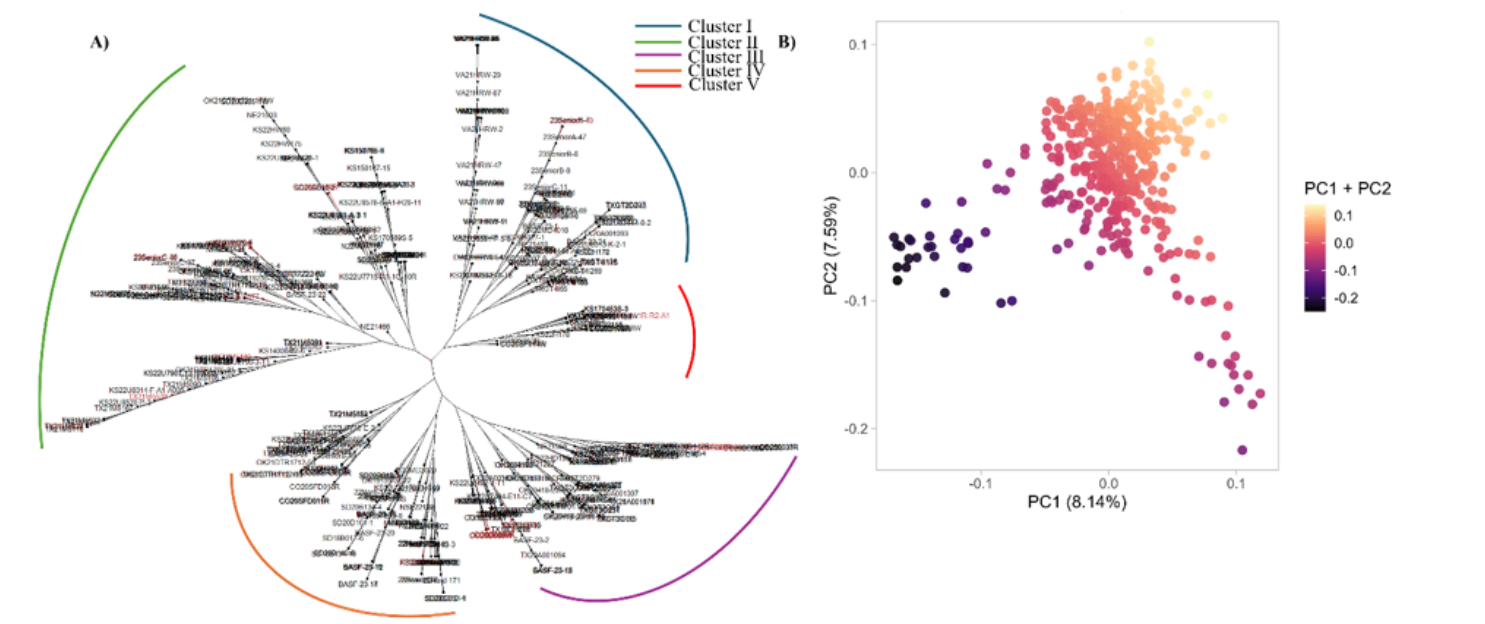


Figure 4

Population structure and genetic diversity analysis of 412 wheat genotypes. **(A)** Phylogenetic tree illustrating the genetic relationships among wheat genotypes, grouped into five major clusters indicated by different colors. **(B)** Three-dimensional principal component analysis (PCA) plot showing the genetic distribution pattern of the genotypes based on the first three principal components (PC1 = 8.14%, PC2 = 7.59 of total variation). Lines highlighted in red indicate resistant genotypes.

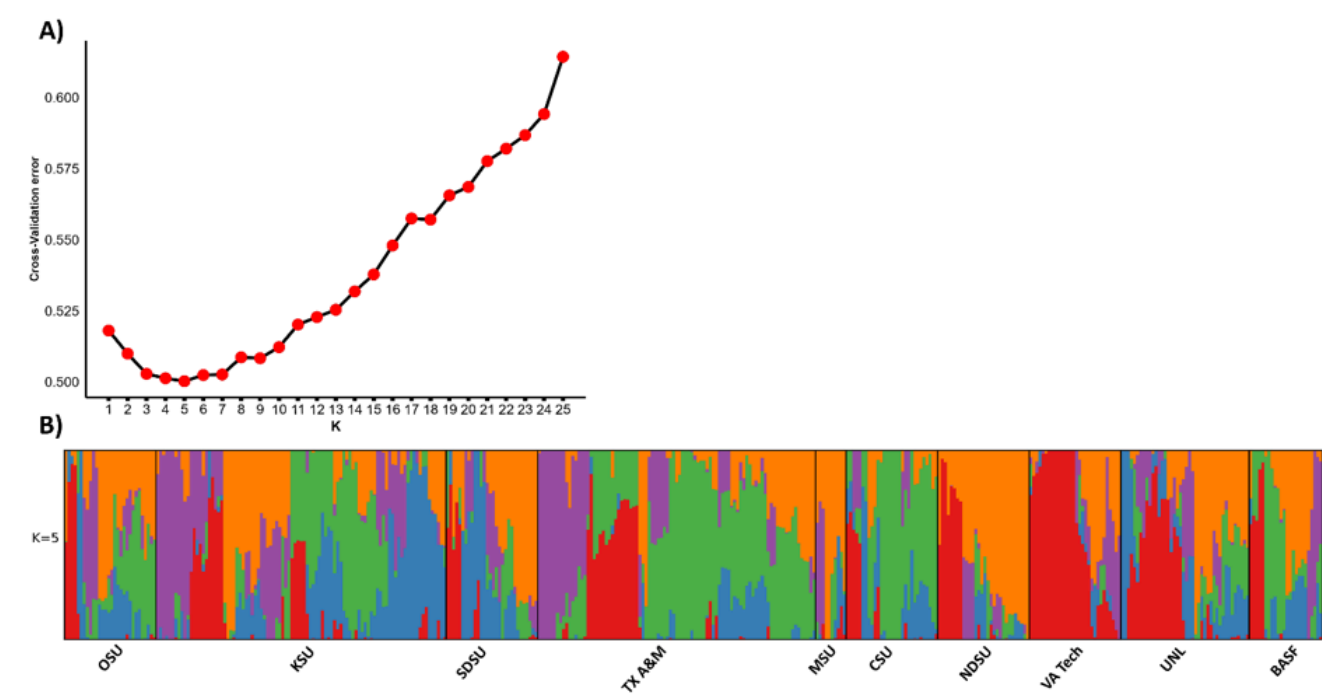


Figure 5

Population structure analysis of RGON wheat accessions using ADMIXTURE.

The upper plot shows the cross-validation (CV) error values across different numbers of ancestral populations (K = 1–25), indicating the lowest CV error at K = 5, which was considered the optimal number of subpopulations. The lower bar plot represents the population structure at K = 5, where each vertical bar corresponds to an individual accession, and colors represent the proportion of inferred genetic ancestry from each of the five subpopulations. Accessions are grouped by their geographical origin or breeding program (Oklahoma State University, Kansas State University, South Dakota State University, Texas A&M, Montana State University, Colorado State University, North Dakota State University, Virginia Tech, University of Nebraska Lincoln, and BASF).

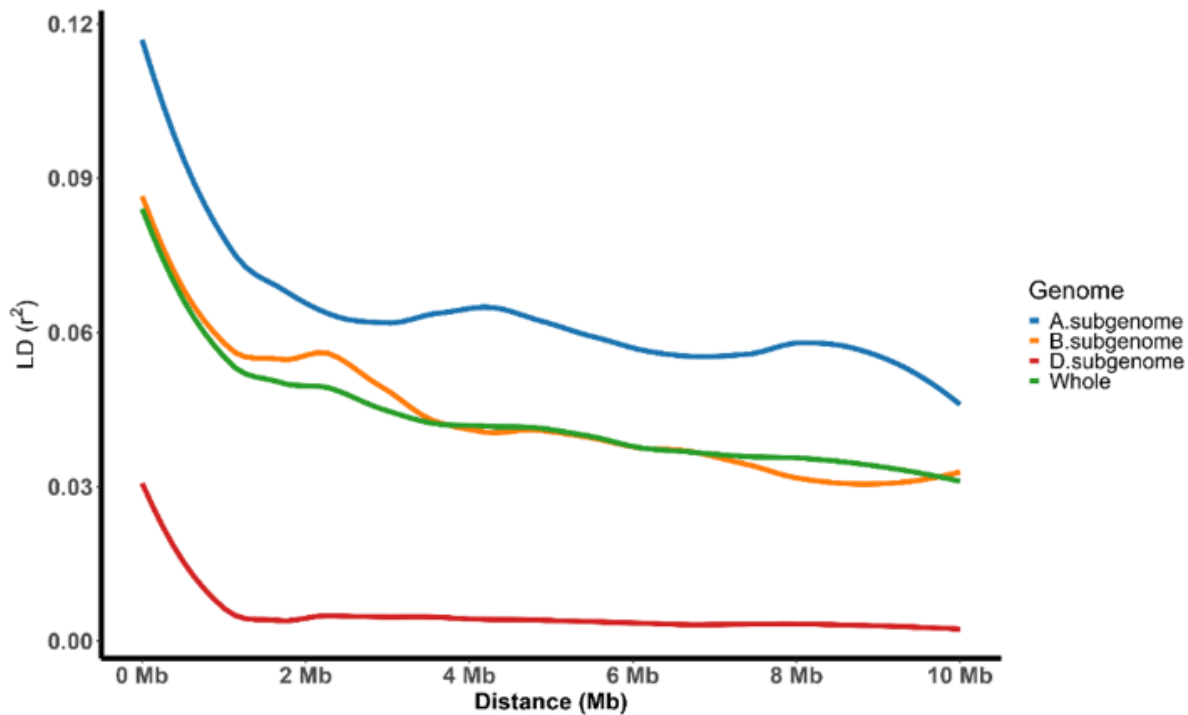


Figure 6

Linkage disequilibrium (LD) decay across the A, B, and D sub genomes and the whole genome of hexaploid wheat in the population of 412 hard winter wheat genotypes and based on 15368 SNP markers. The LD decay (measured as the squared correlation coefficient, r^2) is plotted against the physical distance between SNP pairs (in megabases, Mb). The A sub-genome (blue line) exhibits the highest LD and the slowest decay, followed by the B sub-genome (orange line), while the D sub-genome (red line) shows the lowest LD and the fastest decay. The whole-genome LD decay (green line) represents the average trend across all three sub-genomes. Overall, LD decreases with increasing physical distance, consistent with recombination effects and genomic diversity among sub-genomes.

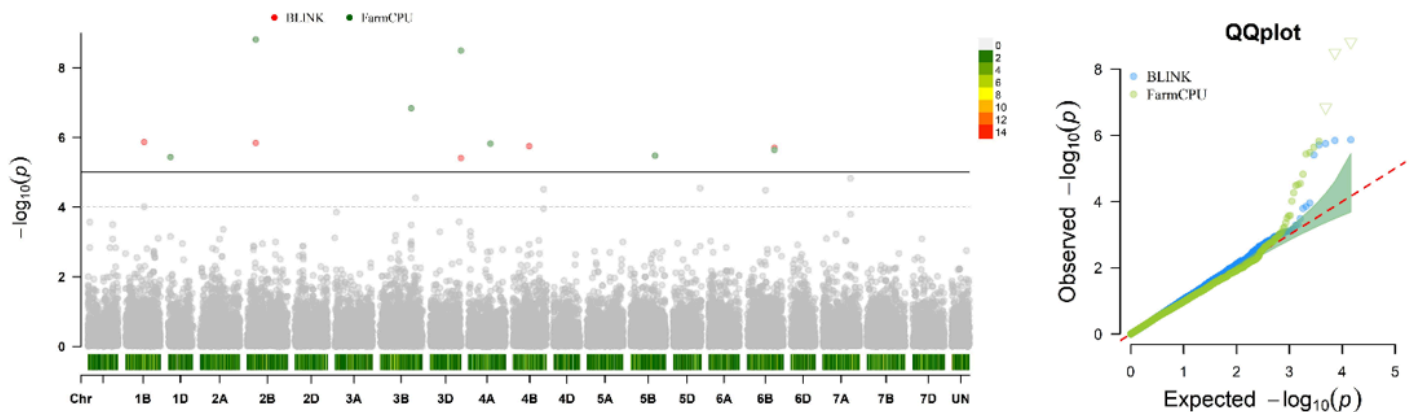


Figure 7

Manhattan and Q-Q plot for SNPs associated with BLS resistance in four GWAS models. Red dots show the significant marker obtain from BLINK while green dots showing the significant marker from FarmCPU. Green dotted line showing $p < 0.05$ and black dotted line showing $p < 0.01$.

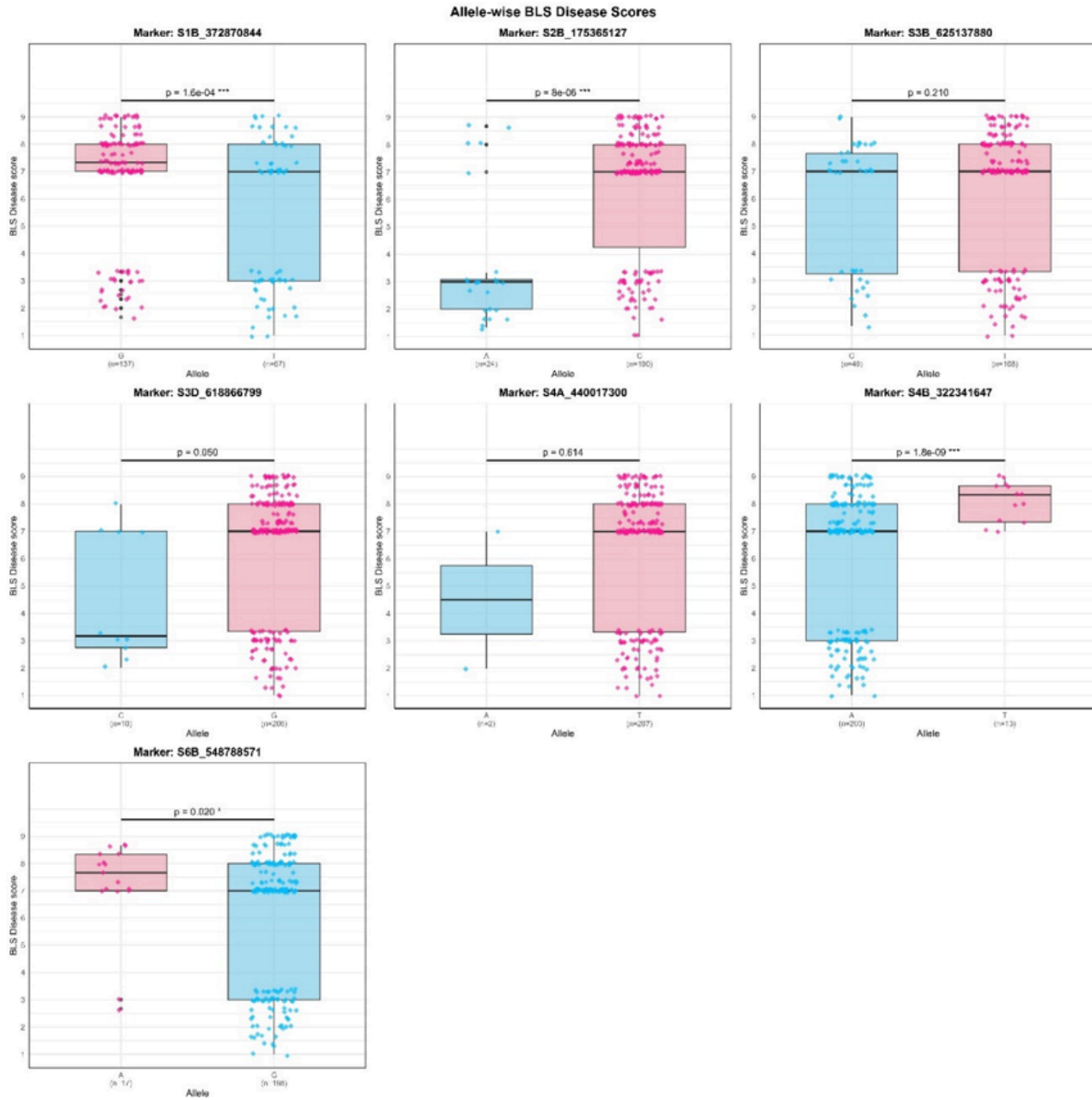


Figure 8

Allelic effects of significant marker–trait associations (MTAs) for BLS response. Allelic effects were estimated using contingency tables. Blue bars represent the proportion of lines with BLS scores of 1–3, while red bars represent the proportion of lines with BLS scores of 7–9. Lines with intermediate BLS scores (4–6) were excluded to simplify the analysis. The x-axis shows the marker alleles, and bar heights indicate the proportion of lines carrying each allele

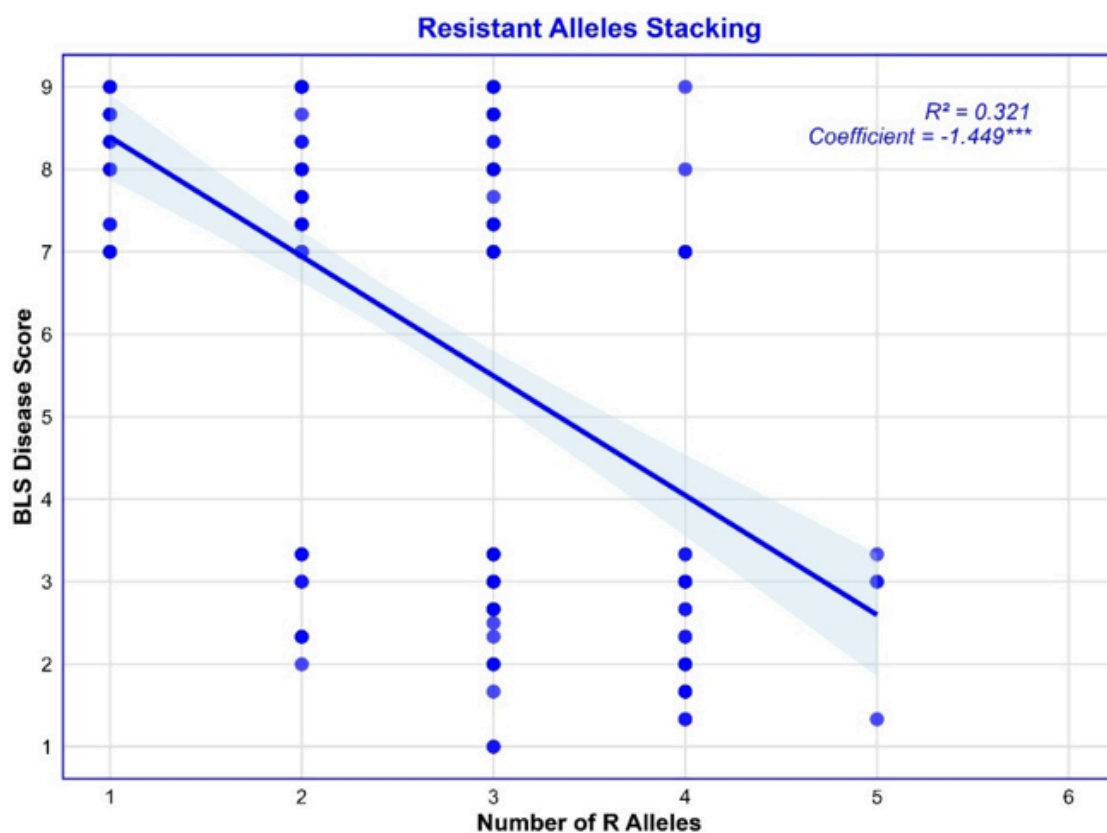


Figure 9

Additive effects of identified marker–trait associations (MTAs) on BLS disease severity. The x-axis depicts the number of Susceptible Loci and y-axis shows BLS disease severity on a scale of 1 to 9, where 1 is highly resistant and 9 is susceptible. The number of resistant (R) alleles at the identified MTAs showed clear additive effects on BLS disease score. Increasing R-allele stacking was associated with reduced disease severity, resulted in lower BLS scores. Lines indicate linear regression fit with 95% confidence intervals.

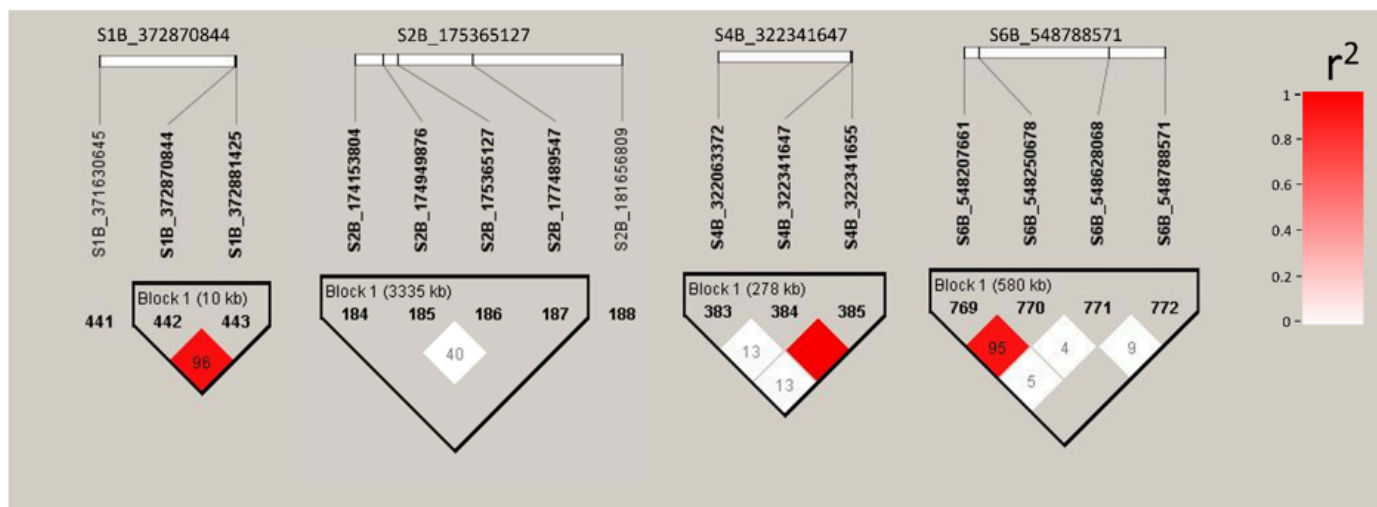


Figure 10

Linkage disequilibrium (LD) blocks surrounding the significantly associated marker–trait associations (MTAs). The color gradient from red to white represents the strength of LD (r^2), where red indicates high LD and white indicates no LD. Inverted triangular shapes denote the observed LD blocks. LD was analyzed by first calculating LD between each MTA and its closest genetic marker. If low LD ($r^2 < 0.4$) was detected, the next closest marker was evaluated. This process was repeated by sequentially including additional neighboring markers until no LD was observed around the MTA.

Supplementary Files

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

- [Supplementaryfile1finalfinal.docx](#)