Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic mapping of a major-effect locus and candidate genes conferring resistance to fungal pathogen Septoria cannabis in Cannabis sativa L.

Jocelyn A. Schwartz
Cornell University
Michael A. Quade
Cornell University
Luis A. Monserrate
Cornell University
Lawrence B. Smart
Cornell University
Christine D. Smart
cds14@cornell.edu
Cornell University
Cornell University

Research Article

Keywords: Hemp, Cannabis sativa, disease, Septoria, NBS-LRR

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11557771/v1

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

Additional Declarations: No competing interests reported.

Title: Genetic mapping of a major-effect locus and candidate genes conferring resistance to fungal pathogen *Septoria cannabis* in *Cannabis sativa* L.

Authors: Jocelyn A. Schwartz[1], Michael A. Quade[2], Luis A. Monserrate[2], Lawrence B. Smart[2], and Christine D. Smart[1,†]

Affiliations:[1]Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech, Geneva, NY 14456 USA.; [2]Horticulture Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech, Geneva, NY 14456 USA.

[†]Corresponding author: Christine D. Smart, cds14@cornell.edu, 315-787-2441

ORCID: Jocelyn A. Schwartz, 0000-0001-7590-4897; Michael A. Quade, 0009-0002-8395-8400; Luis A. Monserrate, 0000-0002-0518-5302; Lawrence B. Smart, 0000-0002-7812-7736; Christine D. Smart, 0000-0003-4162-9502

Key Message: We identified a 724-kb locus involved in resistance to hemp pathogen *Septoria cannabis* through QTL mapping, establishing resources for marker-assisted selection and breeding.

Acknowledgements: We gratefully acknowledge support from Drs. Tyler Gordon and Zachary Stansell for generously sharing hemp germplasm from the USDA-ARS collection; the USDA-ARS PGRU hemp team and the Cornell NE9 farm team for maintaining the field for the mapping population screening; fantastic technical support from the Chris Smart and Larry Smart Labs, in particular Colin Day and Jenna Hershberger; and CannSelect/VCF filtering improvements from the amazing Dr. George Stack.

Abstract

Septoria leaf spot, caused by *Septoria cannabis*, is a common disease of hemp (*Cannabis sativa* L.), yet the genetic basis of resistance to this pathogen is unknown. The objective of this study was to characterize resistance derived from hemp germplasm accession Iowa Feral #20 (G 33241) and identify genomic regions associated with resistance. A resistant F_1 individual derived from a cross between susceptible hemp cultivar 'FL 58' and resistant Iowa Feral #20 (designated GVA-H-25-1515) was self-pollinated to generate an F_2 mapping population. Among 310 F_2 progeny evaluated under field conditions, 236 were resistant and 74 were susceptible, a segregation ratio consistent with a single, major, dominant resistance locus. Bulked-segregant analysis identified a single resistance-associated region on chromosome 3 (chr03), which was independently confirmed by linkage mapping. A major chr03 quantitative trait locus with a maximum logarithm of the odds of 49.86 explained 64.7% of the phenotypic variation in disease severity, as measured using area under the disease progress curve. The initial 1.78-Mb locus was further delimited using molecular markers to a 724 kb interval. This interval contains 11 genes with putative roles in plant defense, including a cluster of seven nucleotide-binding leucine-rich repeat genes. Molecular markers within the resistance-associated locus showed high genotype–phenotype concordance within the mapping population and provide tools for tracking resistance derived from Iowa Feral #20. These results identify the first locus associated with resistance to *S. cannabis* in *C. sativa*, provide candidate genes for functional characterization, and establish genetic resources for breeding *S. cannabis*-resistant hemp.

Keywords

Hemp, *Cannabis sativa*, disease, Septoria, NBS-LRR

Introduction

Septoria leaf spot, caused by the fungal pathogen *Septoria cannabis*, is among the most common diseases of field-grown hemp (*Cannabis sativa* L.) in the northeastern United States and has been reported across numerous *C. sativa*-producing regions worldwide (McPartland et al. 2000; Munir et al. 2023; Thiessen et al. 2020). Disease symptoms manifest as small chlorotic lesions that develop necrotic centers and coalesce as disease progresses, resulting in extensive leaf chlorosis and premature defoliation (Rahnama et al. 2021). Management currently relies primarily on cultural practices that reduce inoculum and leaf wetness, including crop rotation, strip tillage, sanitation, increased plant spacing, and avoidance of overhead irrigation (Gauthier and Smith 2026). Field management with fungicides currently labeled for hemp has had limited success (Schwartz et al. 2025a; Schwartz et al. 2025b). Therefore, the development of resistant *C. sativa* cultivars represents an important complementary strategy for sustainable management of hemp Septoria leaf spot.

Many plant resistance (*R*) genes encode nucleotide-binding leucine-rich repeat (NBS-LRR) proteins that recognize pathogen effectors and activate effector-triggered immunity (ETI), which is often associated with a localized hypersensitive response (Jones and Dangl 2006). Resistance mediated by major-effect *R* genes can produce qualitative, race-specific immune responses, although this may impose selection for pathogen genotypes capable of overcoming host resistance (Milgroom 2017). Resistance-associated quantitative trait loci (QTL) have been mapped in several *Septoria* pathosystems, including diseases of wheat and other crop species (Joshi 2015; Saintenac et al. 2018; Langlands-Perry et al. 2022; Zhong et al. 2006). These studies demonstrate that host resistance to *Septoria* spp. can range from major-effect loci to quantitative

multi-gene resistance and provide a foundation for identifying disease resistance loci in *C. sativa.*

Variation in susceptibility to *S. cannabis* among *C. sativa* genotypes provides an opportunity to develop host resistance as a component of disease management. Recent evaluation of diverse *C. sativa* germplasm under field conditions identified substantial variation in Septoria leaf spot severity, including accessions that consistently exhibited low disease severity across years (Schwartz et al. 2026). These findings demonstrated that sources of resistance are present within available *C. sativa* germplasm; however, the genetic basis of this resistance remains unknown. Identifying loci underlying resistance and developing associated molecular markers are critical steps for incorporating resistance into adapted cultivars through marker-assisted selection.

Despite increasing interest in disease resistant *C. sativa* cultivars, relatively few resistance loci have been genetically characterized (Sirangelo et al. 2023). Most work has focused on hemp powdery mildew, caused by *Golovinomyces ambrosiae* (Roy et al. 2026; Vignale et al. 2026). The dominantly inherited *PM1* and *PM2* loci are localized to regions containing several candidate resistance genes (Mihalyov and Garfinkel 2021; Seifi et al. 2025). Resistance in 'FL 58' is associated with a distinct mechanism involving a loss-of-function mutation in the susceptibility gene *CsMLO1* (Stack et al. 2024). These studies demonstrate that diverse genetic mechanisms for disease resistance occur within *C. sativa* and can be resolved through genetic mapping.

Targeted breeding approaches rely on the knowledge of the genetic basis for economically important traits, which can be identified through QTL mapping studies. Bulked-segregant analysis (BSA), in which individuals from the phenotypic extremes of a segregating population

are pooled and genotyped, can rapidly identify genomic regions associated with major-effect traits while reducing genotyping requirements (Michelmore et al. 1991). Candidate regions identified through BSA can subsequently be resolved and validated through linkage mapping of individual progeny. Markers linked to resistance loci can then facilitate marker-assisted selection, allowing breeders to select resistant individuals without relying exclusively on repeated disease phenotyping. Such tools could accelerate the introgression of Septoria leaf spot resistance into *C. sativa* germplasm.

The objective of this study was to determine the genetic basis of Septoria leaf spot resistance in *C. sativa* and develop genomic tools to facilitate breeding for disease resistance. A biparental F_2 mapping population derived from susceptible 'FL 58' and resistant accession Iowa Feral #20 (U.S. National Plant Germplasm System [NPGS] Accession G 33241) was phenotyped for resistance to *S. cannabis* under field conditions. Specifically, we aimed to (i) characterize the inheritance of Septoria leaf spot resistance, (ii) identify genomic regions associated with resistance using BSA and linkage mapping, (iii) develop molecular markers linked to the resistance locus for marker-assisted selection, and (iv) identify candidate resistance-associated genes within the mapped interval. This work provides the first genetic mapping of resistance to *S. cannabis* in *C. sativa* and establishes a foundation for the introgression of Septoria leaf spot resistance into regionally adapted germplasm.

Materials and Methods

Population Development

Cultivar ‘FL 58’ (Sunrise Genetics, CO) and a selection from accession Iowa Feral #20 (NPGS Accession G 33241; GVA-H-25-1515) were crossed in 2025 to generate the F_1 family designated GVA-H-25-1001. To identify a chemotype III, XY male Iowa Feral #20 plant for the cross, 90 seeds were sown into SureRoots 50 deep-well 50-cell flats (T.O. Plastics, Clearwater, MN) with LM-3 All Purpose Peat Growing Mix (Lambert, Rivière-Ouelle, Québec, Canada) in the Cornell AgriTech greenhouse. Leaf tissue was removed from the newest, fully expanded leaf from 3-week-old plants, placed in a 96-well microplate with one plant sample per well and dried on silica gel. DNA was extracted from the leaf tissue sample plate using a CTAB-based method adapted from Doyle and Doyle (1987). PACE assays were used to determine CSP-1 and CPP-1 (Toth et. al. 2020). The most vigorous chemotype III, male Iowa Feral #20 plant, GVA-H-25-1515, was selected to serve as the pollen parent in a cross with a clonal individual of the cultivar ‘FL 58’ to generate an F_1 family (Supplementary Figure S1).

Two ramets of ‘FL 58’ were propagated from stem cuttings and grown to approximately 1 m in height in the greenhouse under a 16-h photoperiod along with GVA-H-25-1515. Plants were later transferred to a growth chamber with a 12-h photoperiod to induce flowering and prevent pollen contamination. After the GVA-H-25-1515 produced male flowers, a pivoting fan was added to the growth chamber to facilitate pollination of the female ‘FL 58’ clones. After the seed on ‘FL 58’ reached maturity, the plants were dried, threshed, and seeds were cleaned and stored in a freezer (−18 °C) to ensure seed longevity. This F_1 seed lot was designated GVA-H-25-1001.

Twenty seedlings from F_1 family GVA-H-25-1001 were inoculated with *S. cannabis* isolate 22383 (GenBank #PV774436) in a walk-in growth chamber above 80% humidity according to

Schwartz et al. (2025a) at a concentration of 5.76 *10^5 conidia mL^{-1}. The plants were evaluated for disease incidence at 14 days post inoculation, and a female plant that did not display *S. cannabis* symptoms was selected and designated GVA-H-25-1001-007, to generate an F_2 family.

Four ramets of GVA-H-25-1001-007 were propagated by stem cuttings and grown to 1 m in height as described above and transferred to the growth chamber. Silver thiosulfate (STS) was applied to two of the four plants, designated as pollen parents, to force the genetically female plants to produce staminate pollen-bearing flowers (Lubell and Brand 2018). Four foliar spray applications of 8 mM STS every 5 days were sufficient to initiate productive staminate flowers. Once the pollen parents produced staminate flowers, they were planted alongside a GVA-H-25-1001-007 seed parent in a 75.7 L pot containing a tomato cage and covered in CannaLock pollination bags (MidCo, St. Louis, MO) that could be cinched at the base of the pot to prevent pollen contamination. The plants were irrigated using the Raindrip R560DP Drip Watering System (Raindrip, Woodland Hills, CA). After the seed reached maturity, plants were dried, threshed, and seeds were cleaned and designated as family GVA-H-25-1518.

Field phenotyping for resistance to Septoria leaf spot

Seeds from F_2 family GVA-H-25-1518 were sown into SureRoots deep-well 50-cell flats containing LM-3 All Purpose Peat Growing Mix potting mix in the greenhouse on 5 May 2025, where they were regularly maintained and watered. The plants were fertilized with 75 ppm Jack's Professional LX 15-5-15 (J.R. Peters, Inc., Allentown, PA) with 4% Ca and 2% Mg on 15 and 21 May 2025. Raised beds that were 1.2 m wide, with T-Tape drip line irrigation tape set down the center of the beds (Rivulis, San Diego, CA) and covered with 1.25-mm black polyethylene mulch were prepared at the Wellington research farm at Cornell AgriTech on 20 May 2025. The

153 field was fertilized with 22-0-0 ‘ammonium nitrate cut’ at 504.4 kg ha^{-1} (11 kg N ha^{-1}), speed
154 tilled, and clover seeded in the alleys on 19 May 2025. The 310 transplants from F_2 family GVA-
155 H-25-1518 were set into the field by hand on 28 May 2025. Additionally, 10 replicated single-
156 plant plots of each of the parental genotypes, including clones of ‘FL 58’, GVA-H-25-1515, and
157 GVA-H-25-1001-007, were transplanted into the field to serve as controls. Planet Marvel Bloom
158 and Flower Plant Food Nutriculture 12-31-14 (Plant Marvel, Chicago Heights, IL) were added at
159 100 ppm to the transplants. Clover was reseeded between the raised beds on 4 June 2025 to
160 suppress weeds.

161 The field was inoculated with *S. cannabis* isolate 22383 according to Schwartz et al. (2025a) at a
162 concentration of 1.5 $*10^5$ conidia mL^{-1} on 1 July 2025, 5.64 $*10^5$ conidia mL^{-1} on 8 July 2025,
163 1.75 $*10^5$ conidia mL^{-1} on 25 July 2025, and 2.75 $* 10^5$ conidia mL^{-1} on 13 August 2025.
164 Approximately 15 mL of inoculum was applied to each plant at each inoculation. Due to dry
165 weather conditions, water was applied to each plant using a battery-powered backpack sprayer in
166 the mornings from 14 August 2025 to 19 August 2025 to increase leaf wetness conducive for
167 disease development.

168 Disease severity ratings were visually estimated on a scale of 0 to 100% coverage of the adaxial
169 leaf surface of each plant. Disease ratings were completed on 13 August, 3 September, 15
170 September, 22 September, and 29 September 2025. Five disease severity ratings were used to
171 calculate the area under disease progress curve (AUDPC) according to the formula in Jeger and
172 Viljanen-Rollinson (2001). All calculations and analyses of the field phenotyping data were
173 performed in RStudio (Posit Team 2024). The R package *agricolae* (de Mendiburu 2023) was
174 used to calculate AUDPC. Additionally, F_2 individuals were classified as either resistant (disease
175 severity $\leq 2\%$) or susceptible (disease severity $> 2\%$) based on the final disease severity rating on

29 September 2025. Chi-square goodness-of-fit tests were used to determine whether the observed proportions of susceptible and resistant individuals in the F_2 population statistically deviated from a 3:1 ratio expected for a single major-effect dominant locus. Leaf tissue was collected from each plant in the field (n = 341) on 6 October 2025 and immediately dried on silica gel.

Bulked Segregant Analysis

DNA was extracted from each F_2 plant and four replicates of each parent genotype using a modified CTAB protocol (Doyle and Doyle 1987), with DINOMAG ™ SPRI Magnetic Beads (Redoxica, Little Rock, AR) to reversibly bind the DNA, using an Agilent Bravo automated liquid handling platform (Agilent). Bulks for BSA were prepared by pooling equimolar amounts of DNA from each pool representing 30 resistant or 30 susceptible F_2 individuals. Illumina TruSeq libraries were prepared from each pool and sequenced on the NovaSeq X platform at the Cornell Biotechnology Research Center.

Paired-end reads were quality filtered and trimmed using fastp (Chen et al. 2018) and aligned to the ‘FL 58’ haplotype 2 reference genome assembly (Stack et al. 2025) using the NVIDIA Parabricks v4.7.1 fq2bam tool (Zhu et al. 2025), a GPU-accelerated wrapper for running a Burrows-Wheeler Aligner (BWA-MEM; Li and Durbin, 2009). Following this, variants were called using Parabricks haplotypecaller tool, which is a GPU-accelerated wrapper running GATK’s HaplotypeCaller tool that produces a GVCF variant file per bulk. Called variants were then combined and jointly called between bulks using CombineGVCFs and GenotypeGVCFs tools from GATK (Van der Auwera and O'Connor, 2020). The resulting VCF file was filtered

198 with a minimum read depth of 10× and converted to a table format using VCFtools (Danecek et
199 al. 2011). The variant table was further converted to CSV format, and BSA was performed using
200 a modified, parallelized implementation of PyBSASeq (Zhang and Panthee 2020) following the
201 pipeline in Stack et al. (2025). Variants with genotype quality <20 or unusually high read depth
202 (>3 standard deviations above the mean) were excluded. For each retained variant, allele
203 frequencies were calculated for each bulk, and allele-frequency differences (ΔAF), Fisher's exact
204 test *P*-values, and G-statistics were calculated. Variants with Fisher's exact test $P < 0.01$ were
205 classified as significant structural variants (sSV). Genome-wide patterns were evaluated using 2-
206 Mb sliding windows with 10-kb increments. For each window, the proportion of sSV, G-statistic,
207 and ΔAF were calculated. Genome-wide significance thresholds were estimated from 10,000
208 simulations under the null hypothesis of no association, with the 99.5th percentile used as the
209 significance threshold. Candidate peaks exceeding the genome-wide threshold were subsequently
210 evaluated using window-specific simulations.

211

212 **Linkage Mapping**

213 Illumina-compatible DNA libraries from the complete F_2 population, parents, and F_1 individual
214 were prepped using the SureSelect Max DNA Library Prep Kit (Agilent Technologies, Santa
215 Clara, CA) and subsequently probe-captured using the CannSelect custom Agilent SureSelect
216 target enrichment platform for linkage mapping (Wilkerson et al. 2026). These reduced
217 representation libraries were pooled and sequenced using the Illumina NovaSeq X platform at
218 the Cornell Biotechnology Research Center. Pair-end sequencing reads were processed using
219 fastp (Chen et al. 2018) to extract unique molecular identifiers (UMIs) and detect and trim

adapters with quality Phred ≥ 20 and minimum read length = 60 bp. Trimmed reads were then aligned to the 'FL 58' haplotype 2 reference genome (Stack et al. 2025; NBCI BioProject PRJNA1193891) using the GPU-accelerated BWA-MEM-based fq2bam workflow in NVIDIA Clara Parabricks v4.2.1-1 (Zhu et al. 2025). SAMtools (Li et al. 2009) was used to merge and sort BAM files, after which reads with a mapping quality ≥ 20 were retained to remove PCR duplicates using UMI-tools dedup with the unique method (Smith et al. 2017). Variants of each sample were called independently using GPU-accelerated HaplotypeCaller in NVIDIA Clara Parabricks v4.2.1-1 in gVCF mode. Later, those individual gVCFs were jointly genotyped with GLnexus v1.4.3 (Lin et al. 2018) using its GATK configuration to develop a multisample BCF, which was converted to a compressed VCF using BCFtools and bgzip (Danecek et al. 2021; Li et al. 2009). Lastly, biallelic variants and segregating sites with allele frequencies between 0.01 and 0.99 were retained using BCFtools (Danecek et al. 2021) to finalize the VCF file to be used in subsequent quality filtering. Variants were then filtered based on genotype depth, genotype quality, allele balance, missingness, minor allele frequency, and linkage disequilibrium. Variants with read depth <10×, genotype quality score <20, and heterozygous calls whose allele balance was outside 0.25–0.75 were masked as missing. Samples with >90% missing genotype calls, markers with >5% missing data, and markers with minor allele frequency <0.05 were removed. Markers were subsequently pruned for linkage disequilibrium using PLINK 2.0 (Chang et al. 2015; Purcell and Chang 2026) with a 10-kb window, one-variant step, and r^2 threshold of 0.99. Of the 310 F_2 individuals phenotyped, four were excluded due to unusable genotypic data, leaving a total of 306 F_2 individuals for linkage map construction.

Linkage mapping was performed in R using the qtl v1.70 (Broman et al. 2003) and ASMap v1.0.8 (Taylor and Butler 2017) packages. Markers showing extreme segregation distortion

(threshold= $1*10^{-20}$), greater than 5% missing data, and co-location with other markers were removed prior to map construction using the *pullCross* function. Linkage group clustering and genetic distance estimations were performed using the *mstmap* function of the ASMap package with the Kosambi distance function (Kosambi, 1943) following a *p*-value threshold of 1×10^{-16}. The linkage map was iteratively evaluated using marker crossover profiles and recombination heatmaps. Double crossovers per marker were calculated using the *profileMark* function. Markers associated with >2 double crossovers were removed in an iterative process. Linkage-group assignments and orientations were based on marker physical positions in the 'FL 58' Hap 2 reference genome assembly (Stack et al. 2025; NBCI BioProject PRJNA1193891), and markers with inconsistent physical location based on their genetic location were manually removed.

QTL analysis

Quantitative trait locus analysis was performed in R using the qtl package (Broman et al. 2003). AUDPC values were used as the quantitative phenotype and matched to the 306 F_2 individuals in the final linkage map. Genotype probabilities were calculated at 1-cM intervals with the Kosambi map function using the *calc.genoprob* function. Because the AUDPC followed a zero-inflated continuous distribution, the *scanone* function was used to calculate the log of odds (LOD) scores and the 95% LOD significance threshold, genome-wide QTL analysis using the two-part model and the expectation-maximization (EM) algorithm. For the QTL that exceeded the significance threshold, a 1.5-LOD support interval was determined using the *lodint* function to identify peak and flanking markers. The positions of the flanking markers were used to define

the corresponding physical interval in the 'FL 58' genome assembly. The *makeqtl* and *fitqtl* functions were implemented to fit an additive effects model, from which the phenotypic variance explained per detected QTL was estimated using Haley–Knott regression. Mean AUDPC values for the three genotypic classes at the peak marker were used to calculate additive (a) and dominance (d) effects and the absolute dominance-to-additive effect ratio ($|d/a|$) to characterize QTL gene action (Edwards et al. 1987; Stuber et al. 1987).

Molecular marker development

SNP within the QTL interval were selected for conversion to PCR allele competitive extension (PACE; 3Cr Bioscience, Harlow, UK) markers following Toth et al. (2020), including the peak marker and additional SNP distributed across the interval. Flanking sequences surrounding each target SNP were obtained from the 'FL 58' Hap 2 genome assembly and used to design allele-specific PACE assays. For each SNP, two competing allele-specific primers were designed with the polymorphic nucleotide positioned at the 3′ end and appended with FAM- or HEX-compatible universal tails as detailed in the PACE Genotyping Master Mix User Guide (https://3crbio.com/document/pace-user-guide/). A common primer was designed on the opposite strand to generate a short amplicon encompassing the target SNP. Primer sequences were evaluated for melting temperature, GC content, secondary structure, and potential primer-dimer formation prior to synthesis using Integrated DNA Technologies OligoAnalyzer (https://www.idtdna.com/calc/analyzer). Universal PACE primer tails for HEX (GAAGGTCGGAGTCAACGGATT) and FAM (GAAGGTGACCAAGTTCATGCT) were attached to the 5' end of the competitive primers.

PACE markers were subsequently used to refine the QTL interval based on recombination and genotype–phenotype concordance. Because resistance segregated as a dominant trait, AA and AB genotypes were considered consistent with the resistant phenotype, whereas BB genotypes were considered consistent with the susceptible phenotype. For each marker, predictive accuracy was calculated as the proportion of individuals for which the marker genotype class was concordant with the observed phenotype. Recombinant individuals were then examined to identify transitions between the resistant-associated (AA or AB) and susceptible-associated (BB) genotype classes. Transitions between AA and AB were not used to delimit the resistance-associated interval because both genotypes were associated with resistance. Instead, informative recombination events were those in which a transition between AA/AB and BB resulted in a change from genotype–phenotype concordance to discordance, or vice versa. These recombinant breakpoints, together with marker predictive accuracy, were used to further delimit the resistance-associated region.

To evaluate the distribution of marker alleles associated with resistance in the GVA-H-25-1518 population, sequences encompassing selected SNP markers within the refined QTL interval were queried against available *C. sativa* genome assemblies using BLAST. Exact sequence matches corresponding to the Iowa Feral #20-associated marker allele were recorded for each genome. These comparisons were used to assess the distribution of marker alleles among available *C. sativa* genomes and were not interpreted as evidence of resistance in genomes lacking phenotypic data.

Candidate gene identification

Candidate genes within the fine-mapped QTL interval were identified using gene models predicted from the ‘FL 58’ Hap 2 genome assembly annotated using Helixer v0.3.4 (Holst et al. 2026). Predicted proteins were functionally annotated based on sequence similarity and conserved protein domains with Swiss-Prot and InterPro (UniProt Consortium, 2025; Blum et al. 2025) and genes encoding proteins with putative roles in plant disease resistance or defense signaling were prioritized as candidate genes. Particular attention was given to proteins containing domains characteristic of plant immune receptors, including nucleotide-binding site leucine-rich repeat (NBS-LRR/NLR) proteins.

The region homologous to the resistance locus identified in ‘FL 58’ was examined in the Boone County haplotype BCMb (Lynch et al. 2025). Boone County is a feral *C. sativa* accession from Iowa (GRIN Global Accession G 33332) that may possess similar resistance to *S. cannabis* as Iowa Feral #20, providing a genome for comparison with the susceptible ‘FL 58’ reference. NLR proteins predicted within the homologous BCMb interval were compared with ‘FL 58’ predicted proteins. For BCMb NLRs lacking predicted protein homologs in ‘FL 58’, protein sequences were queried against genomic DNA sequence from the homologous ‘FL 58’ interval using TBLASTN. Alignments across high-scoring segment pairs were evaluated to determine the proportion of each BCMb protein represented by homologous genomic sequence in ‘FL 58’. These comparisons were used to distinguish absence of an ‘FL 58’ predicted protein homolog from absence of homologous genomic sequence and to identify gene candidates exhibiting partial or highly divergent genomic homology.

Results

Septoria cannabis resistance segregates as a single-gene dominant trait

Twenty F_1 progeny derived from a cross between 'FL 58', susceptible to *S. cannabis* and GVA-H-25-1515 (Iowa Feral #20; G 33241), resistant to *S. cannabis* (Supplementary Figure S1), were inoculated with *S. cannabis* in the growth chamber and evaluated for disease incidence (Fig. 1A). Four of the 20 F_1 individuals were found to be susceptible to *S. cannabis*, though disease severity was very low, and one F_1 resistant individual (GVA-H-25-1001-007) was selected for pollination to produce F_2 progeny. A total of 310 F_2 progeny (GVA-H-25-1518) and 30 parent controls were inoculated with *S. cannabis* and evaluated for resistance under field conditions. The mean final disease severity for the susceptible parent ('FL 58'), resistant parent (GVA-H-25-1515), and F_1 parent (GVA-H-25-1001-007) was 13.4%, 0.0%, and 0.1% respectively, while the final disease severity ranged 0.0% to 50% in the F_2 population GVA-H-25-1518 (Supplementary Dataset S1; Fig. 1B).

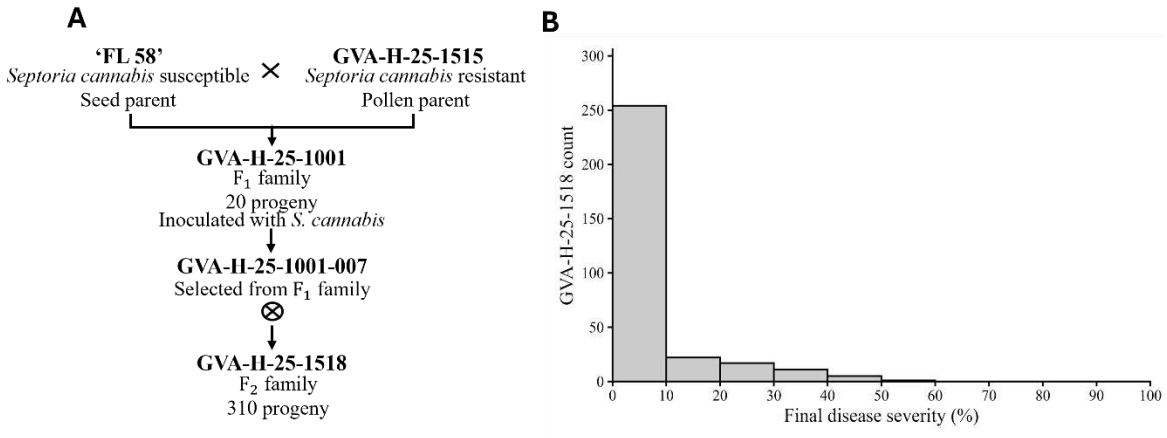


Fig. 1 Development and phenotyping of the GVA-H-25-1518 F_2 population for mapping resistance to *Septoria cannabis*. (A) Population development scheme. (B) Frequency distribution of final *S. cannabis* disease severity among GVA-H-25-1518 F_2 progeny

Using disease severity of ≤ 2% as the threshold for resistance, there were 236 resistant and 74 susceptible progeny. The observed segregation did not differ significantly from the 3:1 resistant-to-susceptible ratio expected for a single-gene dominant resistance locus ($\chi^2 = 0.21$, df = 1, $P = 0.65$; Supplementary Table S1).

BSA identifies a resistance-associated region on chr03

A BSA was performed using DNA from 30 resistant and 30 susceptible F_2 individuals selected from the phenotypic extremes of F_2 segregating population GVA-H-25-1518. A single significant resistance-associated region was mapped to 'FL 58' chr03, where both G-statistic and Δ allele frequency exceeded genome-wide significance thresholds determined from 10,000 simulations (Fig. 2). Nearly all of 'FL 58' chr03 was significant for G-statistic based on the 8.19 G-statistic significance threshold. The majority of a 6.39 Mb region spanning from 75.14 Mb to 81.53 Mb was significant for Δ allele frequency based on the 0.36 confidence interval upper limit. The peak G-statistic and Δ allele frequency were located at 76.36 Mb with the values 23.78 and 0.46. respectively. No significant loci were elsewhere in 'FL 58' genome sequence.

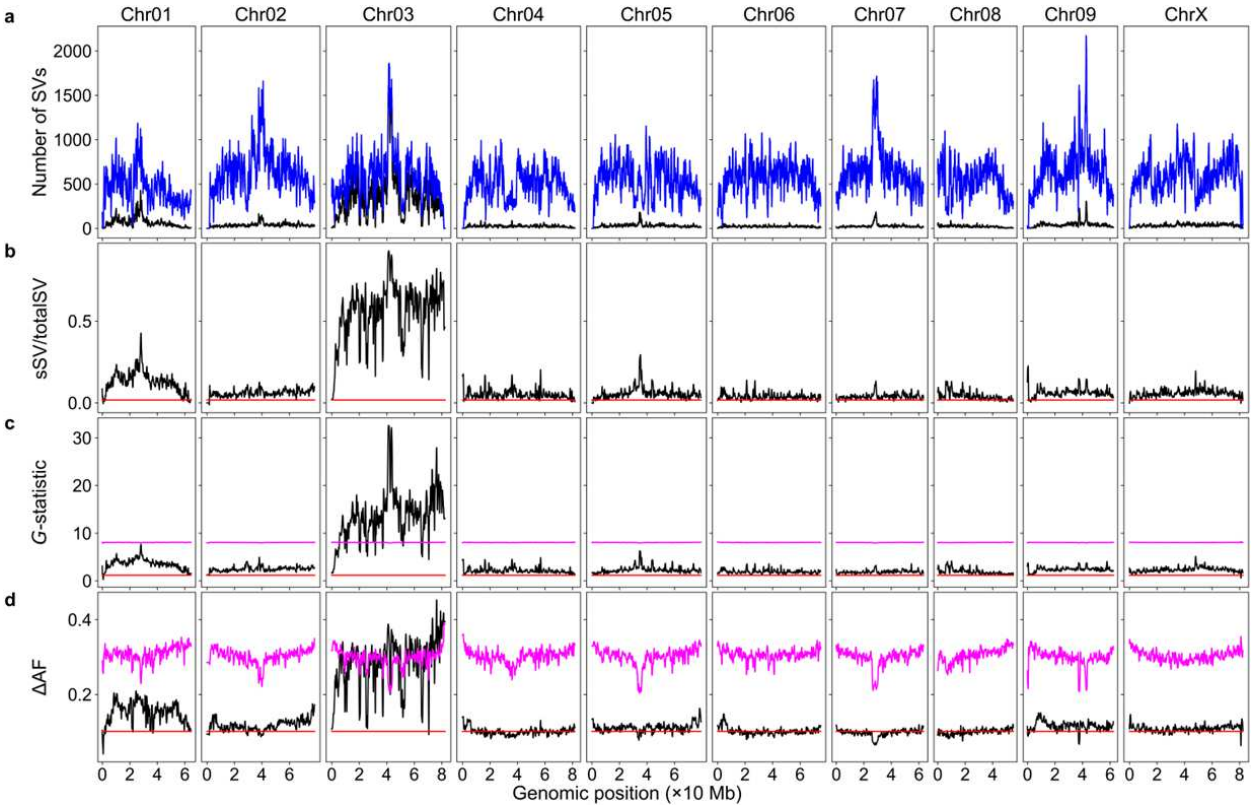


Fig. 2 Bulked segregant analysis of resistance to *Septoria cannabis*. Blue lines indicate structural variant (SV) density and significant SV (sSV) density. Black lines indicate the sSV/totalSV ratio, G-statistic, and Δ allele frequency. Pink horizontal lines indicate genome-wide significance thresholds determined by 10,000 simulations. Horizontal axis indicates genomic position in the 'FL 58' hap 2 genome assembly

Linkage mapping identifies a major-effect QTL for Septoria resistance on chr03

To independently map the resistance locus at the individual-plant level, the F_2 population, parents, and F_1 parent were genotyped using the CannSelect custom Agilent SureSelect target-enrichment platform. Following quality filtering, 1,394 markers and 306 individual F_2 plants were used to construct a 767.8 cM genetic linkage map comprising 10 linkage groups

corresponding to the nine autosomes and X chromosome of *C. sativa* (Fig. 3A; Supplementary Dataset S2).

Genome-wide interval mapping using AUDPC identified a single major-effect QTL on chr03, and no other significant loci in the 'FL 58' genome sequence (Fig. 3B). The QTL peak was at 55.83 cM at marker Chr03_77710713_T_C, which greatly exceeded the genome-wide significance threshold determined by 1,000 permutations with a LOD of 49.86. The 1.5-LOD support interval extended from 55 to 57 cM (Fig. 3C), whose flanking markers delimited a 1.78-Mb physical interval spanning approximately 76.88 to 78.66 Mb on chr03. The locus explained 64.7% of the phenotypic variation in AUDPC. Individuals carrying the resistance-associated allele inherited from Iowa Feral #20 (G 33241) exhibited substantially lower AUDPC than individuals homozygous for the susceptibility-associated allele inherited from 'FL 58'(Fig. 3D). The dominance-to-additive effect ratio was $|d/a| = 0.94$.

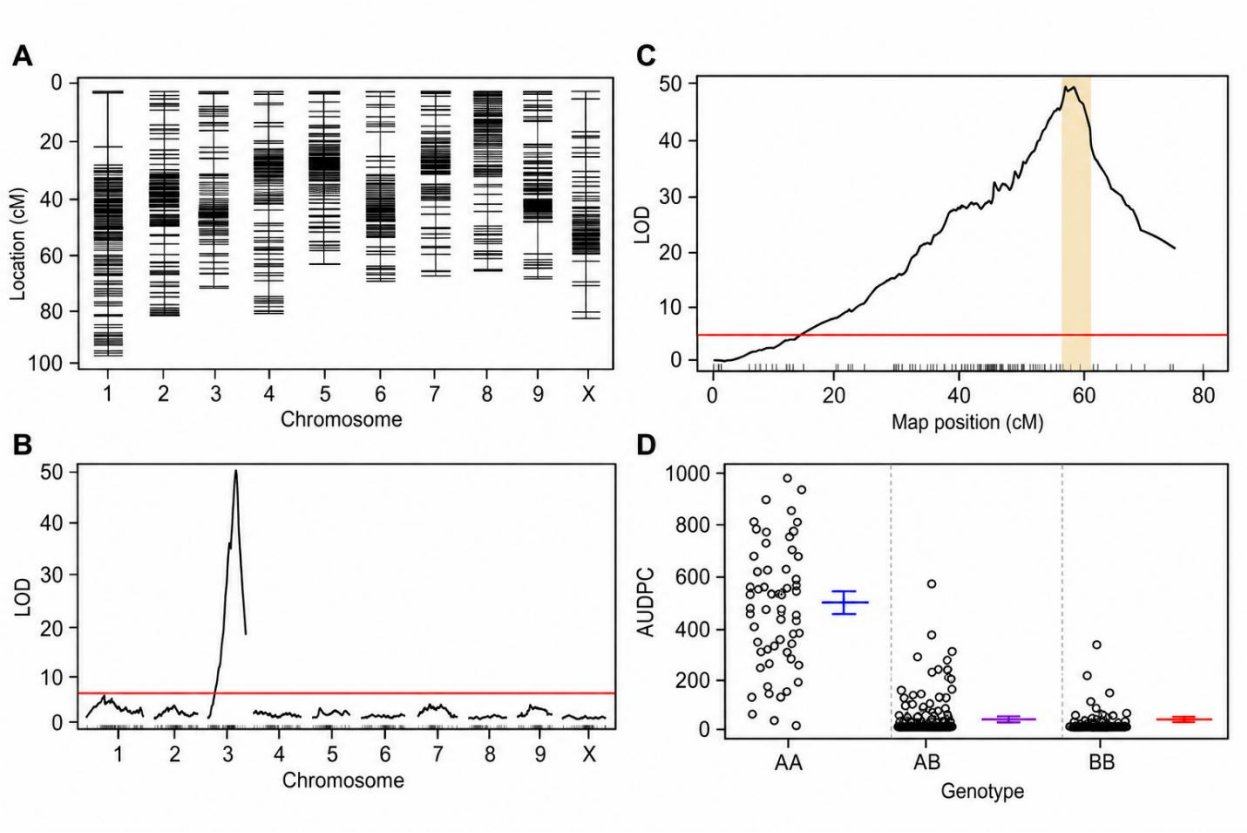


Fig. 3 Linkage map and QTL analysis of Septoria leaf spot resistance in the GVA-H-25-1518 F_2 population. (A) Genetic linkage map showing marker positions across nine autosomes and the X chromosome. Marker positions are shown in centimorgans (cM). (B) Genome-wide QTL mapping for AUDPC. The horizontal red line indicates the genome-wide 95% significance threshold determined by 1,000 permutations. (C) QTL map of AUDPC on chr03. The shaded region indicates the 1.5-LOD support interval surrounding the major QTL peak marker. Tick marks along the x-axis indicate marker positions. (D) Distribution of AUDPC values among individuals grouped by genotype at the peak chr03 marker. Open circles represent individual F_2 plants, and colored bars indicate the mean ± SE for each genotype class. 'A' represents the 'FL 58' allele and 'B' represents the GVA-H-25-1515 allele

PACE markers refine the chr03 resistance locus

Ten PACE markers distributed across the chr03 QTL region were used to genotype F_2 individuals with informative recombinant alleles (Supplementary Table S1, Supplementary Table S2). The PACE assay for the SNP at QTL peak marker Chr03_77710713_T_C was designated SepRes6 (77.71 Mb). Marker–phenotype consistency increased toward the QTL peak, from 75% at SepRes5 (76.82 Mb) to 100% at SepRes12 (77.33 Mb), SepRes6 (77.71 Mb), and SepRes3 (77.82 Mb), before decreasing to 75% at SepRes28 (78.66 Mb).

Recombinant haplotypes further delimited the resistance-associated locus. Susceptible individual P031 remained heterozygous at SepRes11 (77,228,966 bp) but was homozygous for the susceptible-parent allele beginning at SepRes12 (77,334,284 bp), placing the proximal boundary of the resistance locus between SepRes11 and SepRes12. Conversely, resistant individual P009 remained heterozygous through SepRes3 (77,814,557 bp) but was homozygous for the susceptible-parent allele beginning at SepRes15 (77,953,182 bp), placing the distal boundary between SepRes3 and SepRes15 (Supplementary Table S2). These recombination events delimited the resistance-associated region to a maximum interval of 724 kb between SepRes11 and SepRes15, reducing the initial 1.78-Mb QTL interval by approximately 59% (Fig. 4).

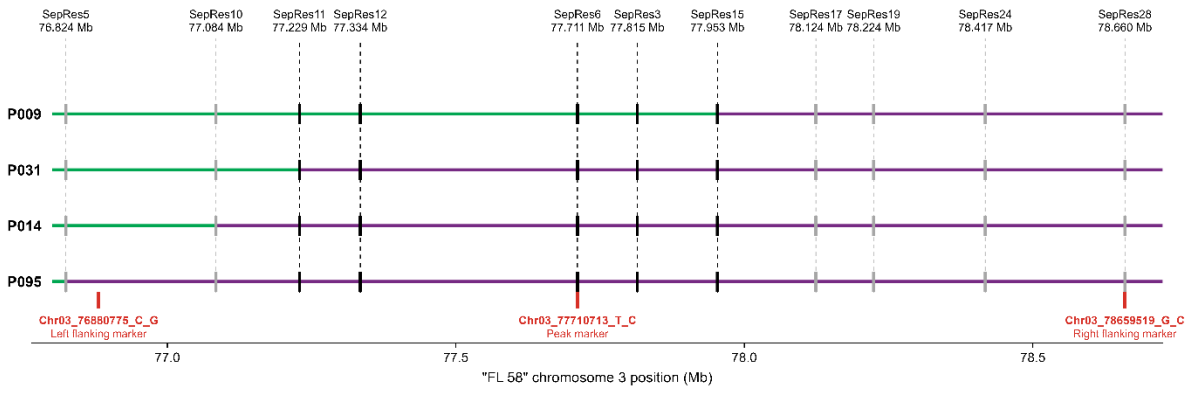


Fig. 4 Refining the major-effect Septoria leaf spot resistance QTL on chr03 of 'FL 58' using PACE markers. Four informative GVA-H-25-1518 F_2 individuals are shown across the QTL interval. Individual P009 were resistant to *S. cannabis* whereas P031, P014, and P095 were susceptible. Green and purple line segments indicate chr03 regions on opposite sides of the inferred recombination breakpoints. Vertical ticks indicate the positions of 11 PACE markers, with markers from SepRes11 through SepRes15, the refined QTL based on recombination events, highlighted in black, and the remaining markers shown in gray. Recombination breakpoints were localized to SepRes15 for P009, SepRes11 for P031, SepRes10 for P014, and SepRes5 for P095. The left and right flanking markers from the CannSelect platform defining the original 1.5-LOD QTL interval and the peak marker are indicated in red. Physical positions correspond to the 'FL 58' reference genome

The refined chr03 resistance interval contains a cluster of NLR genes

Candidate genes within the 724-kb refined locus on chr03 (77.23–77.95 Mb) were identified among Helixer-predicted gene models and functionally annotated using Swiss-Prot protein similarity and InterPro domain predictions. The 724-kb interval contained 62 Helixer-predicted gene models, of which 46 had Swiss-Prot annotations, 51 had InterPro domain annotations, and 11 were identified as putative defense-associated candidates (Supplementary Table S3). Multiple genes within the interval had annotations or predicted protein domains associated with plant defense and immune signaling (Table 1; Supplementary Figure S2). These included genes encoding a leucine-rich repeat receptor-like serine/threonine-protein kinase, an NDR1/HIN1-like protein, NLR proteins, an AP2/ERF-family transcription factor, and a thioredoxin reductase.

439 **Table 1** Putative disease resistance-associated genes within the major QTL interval on chr03 of
440 'FL 58', based on their Swiss-Prot protein annotations and InterPro domain predictions.

Gene model identifier	Position (Mb)	Swiss-Prot annotation	Interpro domain(s)
FL58_000052	77.332–77.336	PXC3-like LRR receptor kinase	LRR; kinase; TM
FL58_000048	77.459–77.460	NDR1/HIN1-like protein 6	NDR1-like; LEA_2
FL58_000046	77.500–77.510	RGA1-like resistance protein	NB-ARC; LRR; CC
FL58_000045	77.514–77.517	RGA1-like resistance protein	NB-ARC; LRR; CC
FL58_000044	77.527–77.531	RGA1-like resistance protein	NB-ARC; LRR
FL58_000040	77.539–77.542	RGA3-like resistance protein	NB-ARC; LRR; CC
FL58_000016	77.599–77.603	RGA1-like resistance protein	NB-ARC; LRR; CC
FL58_000037	77.653–77.656	RGA1-like resistance protein	NB-ARC; LRR
FL58_000036	77.659–77.661	RGA3-like resistance protein	NB-ARC; CC
FL58_000022	77.812–77.816	AP2/ERF transcription factor	AP2/ERF
FL58_000024	77.910–77.916	Thioredoxin reductase 3-like	Thioredoxin reductase

441

442 Seven predicted genes annotated to encode RGA1- or RGA3-like disease resistance proteins
443 occurred in a cluster between approximately 77.50 and 77.66 Mb. All seven contained predicted

NB-ARC domains, and six also contained leucine-rich repeat domains. Predicted coiled-coil domains were identified in several members of the cluster. The NLR cluster was located within the recombinant-defined interval and approximately 50–210 kb upstream of SepRes6 (77,710,713 bp), the marker corresponding to the peak of the chr03 QTL.

Additional defense-associated genes were distributed across the refined interval (Table 1; Supplementary Figure S2). A gene predicted to encode a PXC3-like leucine-rich repeat receptor-like kinase overlapped the position of SepRes12 at approximately 77.33 Mb, while an NDR1/HIN1-like protein was located at approximately 77.46 Mb. Downstream of the NLR cluster, an AP2/ERF-family transcription factor was located at approximately 77.81–77.82 Mb and overlapped SepRes3 (77,814,557 bp). A gene predicted to encode a thioredoxin reductase was located near the distal end of the refined interval at approximately 77.91–77.92 Mb, close to SepRes15.

Iowa Feral #20-associated marker alleles occur in diverse *C. sativa* genomes

To evaluate the broader distribution of SNP alleles associated with resistance in the mapping population, resistance-associated PACE markers within the refined chr03 interval were queried against publicly available *C. sativa* genome assemblies (Supplementary Table S4). Exact matches to the resistance-associated sequences were identified in multiple genetically diverse genomes for each marker examined. The resistance-associated SepRes12 sequence, located within the predicted PXC3-like receptor kinase gene, was particularly widespread among available assemblies, including both feral and cultivated genotypes.

Discussion

This study identified a major-effect locus associated with resistance to *S. cannabis* in *C. sativa* and developed molecular markers predictive of the resistance phenotype. Resistance derived from Iowa Feral #20 (G 33241) segregated in a pattern consistent with a single, dominant resistance gene, and independent BSA and linkage-mapping approaches localized the resistance-associated region to chr03. The chr03 QTL explained 64.7% of the phenotypic variation in AUDPC, and analysis of crossovers in recombinant individuals further reduced the initial 1.78-Mb QTL interval to a 724-kb region. Annotation of this interval identified several candidate genes with potential roles in plant defense, including a cluster of seven *NLR*-like genes. These results provide the first genetic characterization of resistance to *S. cannabis* in *C. sativa* and establish molecular resources for incorporating Septoria leaf spot resistance into adapted germplasm.

Recombinant mapping substantially narrowed the resistance-associated region

The agreement between BSA and linkage mapping provides independent support for the chr03 locus. BSA initially identified a single 6.39-Mb resistance-associated locus on chr03, while linkage mapping using individually genotyped F_2 progeny localized a major QTL to a 1.78-Mb physical interval. Conversion of markers spanning this region to PACE assays subsequently allowed analysis of crossovers in recombinant individuals to refine the locus further. Crossovers in P031 and P009 established opposing boundaries at SepRes11 and SepRes15, respectively, reducing the maximum candidate interval to approximately 724 kb. The marker–phenotype consistency observed across this region provides additional support for localization of the

resistance locus. Fine mapping and/or functional validation strategies will be required to further resolve the locus and identify causal genes.

Compelling candidates for the resistance locus

The concentration of *NLR*-like genes within the recombinant-defined interval provides a plausible biological explanation for the dominant resistance phenotype. Of 62 Helixer-predicted gene models within the 724-kb interval, 11 were prioritized based on annotations or conserved domains associated with plant defense. Most notably, seven *RGA1*- or *RGA3*-like genes formed a cluster between approximately 77.50 and 77.66 Mb. All seven encoded predicted NB-ARC domains, six contained LRR domains, and several also contained predicted coiled-coil domains. NLR proteins function as intracellular immune receptors that recognize pathogen effectors directly or indirectly and activate ETI. *NLR* genes frequently occur in dynamic tandem clusters generated through gene duplication and diversification (Brabham et al. 2025). Major-effect resistance has been characterized in the tomato–*Septoria lycopersci* and the wheat–*Zymoseptoria tritici* pathosystems (Brown et al. 2015; Joshi 2015). QTL associated with resistance to *Septoria tritici* blotch in spring bread wheat included genes encoding NB-ARC domain proteins (Mssilea et al. 2026). The dominant inheritance and large phenotypic effect of the chr03 resistance locus are consistent with an NLR-mediated resistance mechanism. However, the mapping and annotation data alone are insufficient to identify any predicted individual *NLR* coding sequence as the causal resistance gene.

The NLR cluster is not the only plausible immune-receptor candidate within the resistance locus. The refined interval also contains a PXC3-like LRR receptor kinase at approximately 77.33–

77.34 Mb, and marker SepRes12, a SNP located within the predicted gene, showed 100% marker–phenotype concordance in the F_2 individuals evaluated. The presence of this candidate gene is especially noteworthy given that the functionally characterized wheat Septoria resistance genes *Stb6*, *Stb16q*, and *Stb15* all encode receptor-like kinases. *Stb6* encodes a wall-associated receptor kinase-like protein that recognizes the matching *Z. tritici* effector AvrStb6 and confers gene-for-gene resistance (Saintenac et al. 2018), whereas *Stb16q* encodes a cysteine-rich receptor-like kinase that restricts pathogen penetration and early intercellular growth and confers resistance against a broad range of *Z. tritici* isolates (Saintenac et al. 2021). More recently, *Stb15* was shown to encode a G-type lectin receptor-like kinase, further demonstrating the importance and diversity of receptor-like kinases in resistance to *Z. tritici* (Hafeez et al. 2025). Thus, although the cluster of seven *NLR* genes represents an obvious group of candidates based on their predicted immune-receptor functions, an RLK-mediated mechanism analogous to those characterized in the wheat–*Z. tritici* pathosystem cannot be excluded. Additionally, a homolog of this gene was identified in hop (NCBI Reference Sequence: XP_062117722.1). Functional validation and comparison of resistant and susceptible haplotypes will be required to distinguish among these candidate mechanisms.

Several additional genes within the refined interval have annotations consistent with roles in plant defense or stress signaling. The refined interval contains an NDR1/HIN1-like gene, an AP2/ERF-family transcription factor gene, and a predicted thioredoxin reductase gene. These proteins have predicted functions related to extracellular signal perception, defense signaling, transcriptional regulation, and cellular redox responses, respectively. The non-race-specific disease resistance 1/harpin-induced 1-LIKE (*NDR1/HIN1* or *NHL*) gene family is involved in pathogen resistance by inducing CC-NB-LRR activity in *Arabidopsis thaliana* (Century et al.

1995), upregulating salicylic acid (SA) pathways in response to tobacco mosaic virus, *Phytophthora capsici,* and *Pseudomonas syringae* in pepper (Liu et al. 2020), and responding to stressors in other pathosystems (Amato et al. 2025). Additionally, the presence of the late embryogenesis abundant 2 (LEA-2) domain in the 'FL 58' *NDR1/HIN1* gene family is noteworthy, as the LEA-2 domain of wheat *NHL* gene, *TaNHL10*, was shown to interact with effector *ToxA* to facilitate necrosis in the presence of a susceptibility gene (Dagvadorj et al. 2022). The AP2/ERF transcription factor, overlapping marker SepRes3, showed complete genotype–phenotype concordance in the individuals evaluated in the GVA-H-25-1518 F_2 mapping population. The *AP2/ERF* gene family acts downstream of mitogen activated protein kinase (MAPK) cascades and regulates expression of genes associated with hormonal signaling pathways, biosynthesis of secondary metabolites, and formation of physical barriers (Li et al. 2022; Ma et al. 2024; Wang et al. 2020). Thioredoxins mediate reactive oxygen species and are induced in response to infection of cotton and wheat by *Verticillium dahlia and Fusarium graminearum* respectively. (Li et al. 2016; Bhurta et al. 2022). Thus, although the *NLR* cluster represents a compelling group of candidates, further evidence will be required before assigning the resistance function to any particular gene.

The resistance allele may differ structurally from the susceptible 'FL 58' allele

One limitation of candidate-gene identification in this study is that gene prediction was performed using the susceptible 'FL 58' reference genome assembly. Disease-resistance loci, particularly *NLR* clusters, are often among the most structurally variable regions of plant

genomes (Teano et al. 2026), containing high copy number variation at coiled-coil nucleotide-binding site leucine-rich repeat (CNL) clusters (Brabham et al. 2025). Consequently, the gene content and organization of the corresponding interval in resistant Iowa Feral #20 may differ from that in 'FL 58'. The causal resistance allele could represent a divergent sequence of one of the genes annotated in 'FL 58', but it could also involve gene presence/absence variation, copy-number variation (Dolatabadian et al. 2022; Winters et al. 2025) or another structural difference that is not evident from annotation of the susceptible reference alone. The sequence of the resistant haplotype could be obtained by long read sequencing of the F_1 parent (GVA-H-25-1001-007) to generate phased assemblies and would improve the ability compare structural differences at the locus.

The Boone County genome sequence, a feral hemp accession from Iowa, provides additional evidence that the *NLR*-rich region is variable among *C. sativa* haplotypes. Several RPP13-like NLRs predicted in the homologous BCMb interval lacked clear one-to-one predicted protein homologs in 'FL 58' (Supplementary Table S5). BCMb_000021 and BCMb_000022 showed only partial or limited genomic homology to the 'FL 58' interval, whereas BCMb_000020 and BCMb_000024 showed more extensive but divergent homology. These comparisons indicate substantial sequence and/or structural variation within the *NLR*-rich region among *C. sativa* haplotypes. Despite Boone County representing a feral accession from Iowa, we did not phenotype the sequenced Boone County for resistance to *S. cannabis,* thus the relevance of those differences to Iowa Feral #20 resistance remains to be determined.

Linked markers provide tools for breeding for resistance to *S. cannabis*

The widespread occurrence of the resistance-associated SNP across unrelated *C. sativa* genome sequences further indicates that these markers should be interpreted as linked markers rather than causal or universally diagnostic resistance polymorphisms. Their high genotype–phenotype consistency in the GVA-H-25-1518 population reflects linkage to the resistance locus within the Iowa Feral #20-derived haplotype. Once the causal variant is identified by further sequence analysis or functional analysis of gene edited lines, identification of other lines carrying this variant would confirm the distribution of the resistance-associated haplotype within *C. sativa*. Identification of the same resistance haplotype in additional accessions could provide multiple donor sources and help determine whether the locus contributes to resistance across diverse genetic backgrounds.

Beyond identifying the genomic region associated with resistance, the PACE assays developed in this study provide candidate tools for marker-assisted selection tools in introgressing this locus into elite germplasm. SepRes12, SepRes6, and SepRes3 showed complete genotype–phenotype concordance among the individuals evaluated, and all three lie within the recombinant-defined interval. Marker-assisted backcrossing could facilitate the introgression of the *S. cannabis* resistance locus while recovering desirable agronomic characteristics of the recurrent parent. Marker performance should be validated in a targeted elite line prior to initiating introgression of this locus through backcrossing.

Conclusions

This study identified a narrow, major-effect locus on chr03 associated with resistance to *S. cannabis*. The dominant inheritance of resistance, large phenotypic effect of the allele, and

presence of a clustered group of *NLR*-like genes within the recombinant-defined interval provide a strong foundation for identifying the causal resistance gene. Further recombination resolution, comparative analysis of susceptible and resistant haplotypes, and functional characterization of candidate genes will be necessary to determine the molecular basis of resistance. Identification of similar genetic variants across diverse *C. sativa* germplasm and evaluation of the resistance source against multiple *S. cannabis* isolates will additionally determine its utility for breeding durable resistant cultivars.

References

617 Amato V, Mahalath S, Zhang L, Rushton PJ, Shen QJ (2025) Structure and functions of
618 NDR1/HIN1-like (NHL) proteins in plant development and response to environmental stresses.
619 Plant Cell Environ 48:5897–5908. https://doi.org/10.1111/pce.15569

620 Bhurta R, Hurali DT, Tyagi S, Santhee L, Adavi S, Singh D, Mallick N, Chinnusamy V, Vinod,
621 Jha SK (2022) Genome-wide identification and expression analysis of the thioredoxin (*Trx*) gene
622 family reveals its role in leaf rust resistance in wheat (*Triticum aestivum* L.). Front Genet
623 13:836030. https://doi.org/10.3389/fgene.2022.836030

624 Blum M, Andreeva A, Florentino LC et al (2025) InterPro: The protein sequence classification
625 resource in 2025. Nucleic Acids Res 53–D456. https://doi.org/10.1093/nar/gkae1082

626 Brabham HJ, Hernández-Pinzón I, Yanagihara C et al (2025) Discovery of functional NLRs
627 using expression level, high-throughput transformation and large-scale phenotyping. Nat Plants
628 11:2100–2114. https://doi.org/10.1038/s41477-025-02110-w

629 Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses.
630 Bioinformatics 19:889–890. https://doi.org/10.1093/bioinformatics/btg112

631 Brown JKM, Chartrain L, Lasserre-Zuber P, Saintenac C (2015) Genetics of resistance to
632 *Zymoseptoria tritici* and applications to wheat breeding. Fungal Genet Biol 79:33–41.
633 https://doi.org/10.1016/j.fgb.2015.04.017

634 Century KS, Holub EB, Staskawicz BJ (1995) NDR1, a locus of *Arabidopsis thaliana* that is
635 required for disease resistance to both a bacterial and a fungal pathogen. Proc Natl Acad Sci
636 USA 92:6597–6601. https://doi.org/10.1073/pnas.92.14.6597

637 Chang CC, Chow CC, Tellier LC, Shashaank V, Purcell SM, Lee JJ (2015) Second-generation
638 PLINK: Rising to the challenge of larger and richer datasets. Gigascience 4:7.
639 https://doi.org/10.1186/s13742-015-0047-8

640 Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: An ultra-fast all-in-one FASTQ preprocessor.
641 Bioinformatics 34–i890. https://doi.org/10.1093/bioinformatics/bty560

642 Dagvadorj B, Outram MA, Williams SJ, Solomon PS (2022) The necrotrophic effector ToxA
643 from *Parastagonospora nodorum* interacts with wheat NHL proteins to facilitate *Tsn1*-mediated
644 necrosis. Plant J 110:407–418. https://doi.org/10.1111/tpj.15677

645 Danecek P, Auton A, Abecasis G et al (2011) The variant call format and VCFtools.
646 Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330

647 Danecek P, Bonfield JK, Liddle J et al (2021) Twelve years of SAMtools and BCFtools.
648 Gigascience 10. https://doi.org/10.1093/gigascience/giab008

649 de Mendiburu F (2023) agricolae: Statistical procedures for agricultural research. R package
650 version 1.3-7. https://CRAN.R-project.org/package=agricolae

651 Dolatabadian A, Yuan Y, Bayer PE, Petereit J, Severn-Ellis A, Tirnaz S, Patel D, Edwards D,
652 Batley J (2022) Copy number variation among resistance genes analogues in *Brassica napus*.
653 Genes 13:2037. https://doi.org/10.3390/genes13112037

654 Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf
655 tissue. Phytochem Bull 19:11–15

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125. https://doi.org/10.1093/genetics/116.1.113

Gauthier N, Smith H (2026) Integrated disease management in hemp cultivation. In: Noller EJS (ed) Achieving sustainable cultivation of hemp. Burleigh Dodds Science Publishing Limited, Cambridge, UK, pp 279–310

Hafeez AN, Chartrain L, Feng C et al (2025) *Septoria tritici* blotch resistance gene *Stb15* encodes a lectin receptor-like kinase. Nat Plants 11:410–420. https://doi.org/10.1038/s41477-025-01920-2

Holst F, Bolger AM, Kindel F, Günther C, Maß J, Triesch S, Kiel N, Saadat N, Ebenhöh O, Usadel B, Schwacke R, Weber APM, Bolger ME, Denton AK (2026) Helixer: Ab initio prediction of primary eukaryotic gene models combining deep learning and a hidden Markov model. Nat Methods 23:732–739. https://doi.org/10.1038/s41592-025-02939-1

Jeger MJ, Viljanen-Rollinson SLH (2001) The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. Theor Appl Genet 102:32–40. https://doi.org/10.1007/s001220051615

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323–329. https://doi.org/10.1038/nature05286

Joshi BK (2015) Molecular markers for Septoria leaf spot (*Septoria lycopersicii* Speg.) resistance in tomato (*Solanum lycopersicum* L.). Nepal J Biotechnol 3:40–47. https://doi.org/10.3126/njb.v3i1.14230

Kosambi DD (1943) The estimation of map distances from recombination values. Ann Eugen 12:172–175. https://doi.org/10.1111/j.1469-1809.1943.tb02321.x

Langlands-Perry C, Cuenin M, Bergez C, Krima SF, Gelisse S, Sourdille P, Valade R, Marcel TC (2022) Resistance of the wheat cultivar 'Renan' to Septoria leaf blotch explained by a combination of strain specific and strain non-specific QTL mapped on an ultra-dense genetic map. Genes 13:100. https://doi.org/10.3390/genes13010100

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li H, Handsaker B, Wysoker A et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li Y, Liu K, Tong G, Xi C, Jiu J, Wang J, Ren D, Han S (2022) MPK3/MPK6-mediated phosphorylation of ERF72 positively regulates resistance to *Botrytis cinerea* through directly and indirectly activating the transcription of camalexin biosynthesis enzymes. J Exp Bot 73:413–428. https://doi.org/10.1093/jxb/erab415

Li Y-B, Han L-B, Wang H-Y, Zhang J, Sun S-T, Feng D-Q, Yang C-L, Sun Y-D, Zhong N-Q, Xia G-X (2016) The thioredoxin *GbNRX1* plays a crucial role in homeostasis of apoplastic reactive oxygen species in response to *Verticillium dahliae* infection in cotton. Plant Physiol 170:2392–2406. https://doi.org/10.1104/pp.15.01930

Lin MF, Rodeh O, Penn J, Bai X, Reid JG, Krasheninina O, Salerno WJ (2018) GLnexus: Joint variant calling for large cohort sequencing. bioRxiv 343970. https://doi.org/10.1101/343970

Liu C, Peng H, Li X, Liu C, Lv X, Wei X, Zou A, Zhang J, Fan G, Ma G, Ma L, Sun X (2020) Genome-wide analysis of *NDR1/HIN1-like* genes in pepper (*Capsicum annuum* L.) and functional characterization of *CaNHL4* under biotic and abiotic stresses. Hortic Res 7:93. https://doi.org/10.1038/s41438-020-0318-0

Lubell JD, Brand MH (2018) Foliar sprays of silver thiosulfate produce male flowers on female hemp plants. HortTechnology 28:743–747. https://doi.org/10.21273/HORTTECH04188-18

Lynch RC, Padgitt-Cobb LK, Garfinkel AR et al (2025) Domesticated cannabinoid synthases amid a wild mosaic cannabis pangenome. Nature 643:1001–1010. https://doi.org/10.1038/s41586-025-09065-0

Ma N, Sun P, Li Z-Y, Zhang F-J, Wang X-F, You C-X, Zhang C-L, Zhang Z (2024) Plant disease resistance outputs regulated by AP2/ERF transcription factor family. Stress Biol 4:2. https://doi.org/10.1007/s44154-023-00140-y

McPartland JM, Clarke RC, Watson DP (2000) Hemp diseases and pests: Management and biological control—an advanced treatise. CABI, Wallingford

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832. https://doi.org/10.1073/pnas.88.21.9828

Mihalyov PD, Garfinkel AR (2021) Discovery and genetic mapping of *PM1*, a powdery mildew resistance gene in *Cannabis sativa* L. Front Agron 3:720215. https://doi.org/10.3389/fagro.2021.720215

718 Milgroom MG (2017) Durability and the deployment of disease resistance. In: Milgroom MG
719 (ed) Population biology of plant pathogens: Genetics, ecology, and evolution. The American
720 Phytopathological Society, St. Paul, pp 249–273

721 Mssilea FZ, Henkrar F, Lahrichi K, Mrabti SE, Derbounde H, Taesse W (2026) Genome-wide
722 association mapping of adult plant resistance to Septoria tritici blotch in spring bread wheat.
723 Cereal Res Commun 54:1669–1689. https://doi.org/10.1007/s42976-026-00846-5

724 Munir M, Leonberger K, Kesheimer K et al (2023) Occurrence and distribution of common
725 diseases and pests of U.S. cannabis: A survey. Plant Health Prog 24:498–507.
726 https://doi.org/10.1094/PHP-01-23-0004-S

727 Posit Team (2024) RStudio: Integrated development environment for R. Posit Software, PBC,
728 Boston, MA. http://www.posit.co/

729 Purcell S, Chang C (2026) PLINK 2.0, version 2.0. Cognitive Genomics Lab. https://www.cog-
730 genomics.org/plink/2.0/. Accessed 26 Sep 2026

731 Rahnama M, Szarka D, Li H, Dixon E, Castlebury L, Gauthier NA (2021) Reemergence of
732 Septoria leaf spot caused by *Septoria cannabis* on hemp in Kentucky. Plant Dis 105:2286–2289.
733 https://doi.org/10.1094/PDIS-12-20-2620-SC

734 Roy F-A, Asselin Y, Labbé C et al (2026) Hybrid genome assembly of the cannabis powdery
735 mildew agent *Golovinomyces ambrosiae* uncovers important resources for deciphering virulence
736 factors. BMC Genomics 27:532. https://doi.org/10.1186/s12864-026-12995-2

737 Saintenac C, Cambon F, Aouini L et al (2021) A wheat cysteine-rich receptor-like kinase confers
738 broad-spectrum resistance against *Septoria tritici* blotch. Nat Commun 12:433.
739 https://doi.org/10.1038/s41467-020-20685-0

740 Saintenac C, Lee W-S, Cambon F et al (2018) Wheat receptor-kinase-like protein *Stb6* controls
741 gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. Nat Genet 50:368–374.
742 https://doi.org/10.1038/s41588-018-0051-x

743 Schwartz JA, April SG, Day CTC, Smart CD (2025a) Efficacy of organic fungicides for
744 management of hemp Septoria leaf spot. Plant Health Prog 26:602–608.
745 https://doi.org/10.1094/PHP-04-25-0127-RS

746 Schwartz JA, Hershberger JL, Day CTC, Smart CD (2025b) Evaluation of fungicides to manage
747 hemp Septoria leaf spot in Geneva, NY, 2025. Plant Health Prog 26:721.
748 https://doi.org/10.1094/PHP-09-25-0219-PDMR

749 Schwartz JA, Vignale L, Gordon T, Stansell Z, Smart LB, Smart CD (2026) Screening of
750 *Cannabis sativa* germplasm for susceptibility to Septoria leaf spot. Plant Dis.
751 https://doi.org/10.1094/PDIS-07-25-1367-RE

752 Seifi S, Leckie KM, Giles I, O'Brien T, MacKenzie JO, Todesco M, Rieseberg LH, Baute GJ,
753 Celedon JM (2025) Mapping and characterization of a novel powdery mildew resistance locus
754 (*PM2*) in *Cannabis sativa* L. Front Plant Sci 16:1543229.
755 https://doi.org/10.3389/fpls.2025.1543229

756 Sirangelo TM, Ludlow RA, Spadafora ND (2023) Molecular mechanisms underlying potential
757 pathogen resistance in *Cannabis sativa*. Plants 12:2764. https://doi.org/10.3390/plants12152764

Smith T, Heger A, Sudbery I (2017) UMI-tools: Modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. Genome Res 27:491–499. https://doi.org/10.1101/gr.209601.116

Stack GM, Cala AR, Quade MA et al (2024) Genetic mapping, identification, and characterization of a candidate susceptibility gene for powdery mildew in *Cannabis sativa*. Mol Plant Microbe Interact 37:51–61. https://doi.org/10.1094/MPMI-04-23-0043-R

Stack GM, Quade MA, Wilkerson DG et al (2025) Comparison of recombination rate, reference bias, and unique pangenomic haplotypes in *Cannabis sativa* using seven *de novo* genome assemblies. Int J Mol Sci 26:1165. https://doi.org/10.3390/ijms26031165

Stuber CW, Edwards MD, Wendel JF (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci 27:639–648. https://doi.org/10.2135/cropsci1987.0011183X002700040006x

Taylor J, Butler D (2017) R package ASMap: Efficient genetic linkage map construction and diagnosis. J Stat Softw 79:1–29. https://doi.org/10.18637/jss.v079.i06

Teano G, Roudaut A, Gratias A, Geffroy V (2026) Bridging paradoxes in recombination at NLR cluster: A structural genomics perspective. Curr Opin Plant Biol 92:102908. https://doi.org/10.1016/j.pbi.2026.102908

The UniProt Consortium (2025) UniProt: The universal protein knowledgebase in 2025. Nucleic Acids Res 53–D617. https://doi.org/10.1093/nar/gkae1010

Thiessen LD, Schappe T, Cochran S, Hicks K, Post AR (2020) Surveying for potential diseases and abiotic disorders of industrial hemp (*Cannabis sativa*) production. Plant Health Prog 21:321–332. https://doi.org/10.1094/PHP-03-20-0017-RS

Toth JA, Stack GM, Cala AR, Carlson CH, Wilk RL, Crawford JL, Viands DR, Philippe G, Smart CD, Rose JKC, Smart LB (2020) Development and validation of genetic markers for sex and cannabinoid chemotype in *Cannabis sativa* L. GCB Bioenergy 12:213–222. https://doi.org/10.1111/gcbb.12667

Van der Auwera GA, O'Connor BD (2020) Genomics in the cloud: Using Docker, GATK, and WDL in Terra, 1st edn. O'Reilly Media, Sebastopol

Vignale L, Schwartz JA, Stansell Z, Gordon T, Smart CD, Smart LB (2026) Screening diverse *Cannabis sativa* germplasm for resistance to *Golovinomyces ambrosiae*. Plant Dis 110(7). https://doi.org/10.1094/PDIS-07-25-1591-RE

Wang J-H, Gu K-D, Han P-L, Yu J-Q, Wang C-K, Zhang Q-Y, You C-X, Hu D-G, Hao Y-J (2020) Apple ethylene response factor *MdERF11* confers resistance to fungal pathogen *Botryosphaeria dothidea*. Plant Sci 291:110351. https://doi.org/10.1016/j.plantsci.2019.110351

Wilkerson DG, Stack GM, Carlson CH, Quade MA, Dowling CA, Toth JA, Murdock MJ, Jasinski J, Stansell ZJ, McKay JK, Smart LB (2026) CannSelect: A high-quality genotyping platform for *Cannabis sativa*. bioRxiv 2026.08.18.745408. https://doi.org/10.64898/2026.08.18.745408

Winters NP, Wafula EK, Timilsena PR, Ralph PE, Maximova SN, dePamphilis CW, Guiltinan MJ, Marden JH (2025) Local gene duplications drive extensive NLR copy number variation

across multiple genotypes of *Theobroma cacao*. G3 (Bethesda) 15. https://doi.org/10.1093/g3journal/jkaf147

Zhang J, Panthee DR (2020) PyBSASeq: A simple and effective algorithm for bulked segregant analysis with whole-genome sequencing data. BMC Bioinformatics 21:99. https://doi.org/10.1186/s12859-020-3435-8

Zhong S, Toubia-Rahme H, Steffenson BJ, Smith KP (2006) Molecular mapping and marker-assisted selection of genes for Septoria speckled leaf blotch resistance in barley. Phytopathology 96:993–999. https://doi.org/10.1094/PHYTO-96-0993

Zhu T, Vats P, Onken S, Dunstan A, Zamirai B, Puleri DF, Nair A, Oliva M, Gaihre A, Sadhnani P, Li S, Arumugam K, Chacon A, Maric M, Cohen J, Sethia A, Samadi M (2025) Parabricks: GPU accelerated universal pan-instrument genomics analysis software suite. bioRxiv 2025.07.23.666378. https://doi.org/10.1101/2025.07.23.666378

Statements & Declarations

Author contribution statement

Formal analysis, investigation, methodology, and software were performed by J.A.S., M.A.Q., and L.A.M. Conceptualization, data curation, and resources were contributed by J.A.S. and M.A.Q. Project administration, validation, and visualization were contributed by J.A.S. L.B.S. and C.D.S contributed to funding acquisition and supervision. The first draft of the manuscript was written by J.A.S. and all authors commented on previous version of the manuscript. All authors read and approved of the final manuscript.

Funding

This work was supported by the Supplemental and Alternative Crops program, project award no. 2023-38624-40847, from the United States Department of Agriculture's National Institute of Food and Agriculture and New York State Department of Agriculture and Markets contracts CM04068FR and CM04068IR.

Competing Interests

The authors do not declare any financial or non-financial interests.

Data Availability

The selected resistant parent from Iowa Feral #20 (U.S. National Plant Germplasm System [NPGS] Accession G 33241), GVA-H-25-1515, and resistant F_1 parent GVA-H-25-1001-007 can be requested from the corresponding author on reasonable request. The BSA sequence was deposited in NCBI BioProject PRJNA1538723, under BioSamples and SRA accession numbers SAMN63943683 and SRR41002569 for the bulk resistant to *S. cannabis,* and SAMN63943684 and SRR41002568 for the bulk susceptible to *S. cannabis*. Custom scripts used for SNP calling, alignment, filtering, linkage mapping, and QTL analysis are available on the Cornell Hemp GitHub website (https://github.com/CornellHemp/Schwartz_2026_Septoria_Resistance).

Supplementary Files

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

- SupplementaryDatasets.xlsx
- SupplementaryFigures.xlsx
- SupplementaryTables.xlsx