

Linked candidate genes of different functions for white mold resistance in common bean (*Phaseolus vulgaris*. L) are identified by QTL-based pooled sequencing

Atena Oladzadabbasabadi (✉ atena.oladzad@ndsu.edu)

North Dakota State University <https://orcid.org/0000-0001-9455-380X>

Sujan Mamidi

HudsonAlpha Institute for Biotechnology

Phillip N. Miklas

"USDA-ARS Pacific West Area"

Rian Lee

North Dakota State University

Phillip McClean

North Dakota State University

Research article

Keywords: bulk segregant analysis, dry bean, QTL sequencing, *Sclerotinia sclerotiorum*

Posted Date: August 28th, 2019

DOI: <https://doi.org/10.21203/rs.2.13657/v1>

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

Abstract

Background White mold (WM) is a major disease in common bean (*Phaseolus vulgaris* L.), and its complex quantitative genetic control has limited the development of WM resistant cultivars. WM2.2 is one of the nine meta-QTL that has a major effect on WM tolerance. This QTL explains up to 35% of the phenotypic variation and was previously mapped to a large interval on Pv02. Our objective was to narrow the interval of this QTL using QTL-based bulk segregant analysis.

Results The phenotypic and genotypic data from two RIL populations (R31 and Z0726-9), which possess different genetic backgrounds for white mold resistance, were used to select resistant and susceptible lines to generate subpopulations for bulk DNA sequencing, and reads were aligned against the sequence of the resistance parent. The QTL physical intervals for each RIL population were mapped by fixed SNPs in 10kb-2kb sliding windows. WM2.2 QTL was split into two regions WM2.2a (3.54-4.56 Mbp; euchromatic) and WM 2.2b (12.19 to 26.41 Mbp; heterochromatic) in populations R31 and Z0726-9, respectively. For each QTL interval, the possible functional contribution of significant non-synonymous and synonymous polymorphisms was investigated. Gene models encoding for pentatricopeptide repeat, gibberellin 2-oxidase, and heat-shock proteins are the likely candidate genes associated with WM2.2a resistance. A TIR-NBS-LRR class of disease resistance protein and a EF-TU receptor are potential candidate genes associated with WM2.2b resistance and most likely trigger a physiological resistance response to WM.

Conclusion QTL-based pooled sequencing analysis revealed that the large genomic region associated with WM2.2 meta QTL consists of two major QTL each associated with a different WM resistance function. WM2.2a region is most likely associated with avoidance mechanisms while WM2.2b region triggers physiological resistance.

Background

White mold (WM), caused by *Sclerotinia sclerotiorum* (Lib.) de Bary, is a serious worldwide fungal disease in cool, moist conditions. Depending on the common bean genotype, it can cause from 30% (e.g. Central high plains in United States, in 1973) to 100% (e.g., Argentina in 2011) yield losses [1]. All the aerial parts of the plant can be infected by this disease. A white, cottony appearance of mycelium growing on the surface of pods, leaves, branches and stems is a visual characteristic of this disease. As symptoms progress water-soaking followed by desiccation of the affected tissue leads to a dried bleached appearance of stems and branches [1, 2]. Under field conditions, *S. sclerotiorum* acts as a hemi-biotroph, first using biotrophic infection processes and later necrotrophic mechanisms to invade its host. In the biotrophic period, the pathogen uses living tissues of the plant as a source of nutrition and suppresses the host immune system. Eventually, the pathogen kills host tissue extensively and uses the dead cells as nutrients to progress its life-cycle [3]. The long-term survival of *S. sclerotiorum* is mainly via sclerotia (hardened mycelium) formation. These structures are formed on/within infected plants during the necrotrophic period and deposited in the soil, allowing the fungus to survive for years until environmental conditions become favorable [4, 5].

Complete resistance to this disease does not exist in commercial common bean cultivars [6, 7, 8, 9]. Partial resistance, available in some common bean materials, is associated with avoidance or physiological resistance mechanisms, or a combination of both mechanisms. Architectural features such as canopy density, canopy

height, and lodging resistance [5,10] influence avoidance mechanisms. Physiological resistance involves plant defense against pathogen effectors [11].

The complex, quantitative genetic control of both avoidance and physiological mechanisms, along with environmental effects, has limited the development of WM resistance in common bean. Although conventional breeding has provided a limited number of lines with partial resistance, there is a need for more effective tools for routine generation of lines with high levels of WM resistance in common bean. Recently, Vasconcellos et al. [9] presented a meta-analysis that identified 37 QTL associated with partial WM resistance across multiple populations and environments. These QTL, identified using traditional QTL analyses, clustered into nine meta-QTL: WM1.1, WM2.2, WM3.1, WM5.4, WM6.2, WM7.1, WM7.4, WM7.5 and WM8.3. These meta-QTL had significant effects across multiple environments, and the QTL came from different genetic backgrounds. In most cases the QTL region intervals were mega-bases in width. Among these nine meta-QTL, the highest percentage of phenotypic variation was explained by WM2.2 located within the 4.54–22.98 Mb interval on chromosome Pv02 [9]. This QTL was reported in at least six populations and explained up to 35% of the phenotypic variation for disease reaction in both field (avoidance and physiological mechanisms) and greenhouse (physiological mechanisms) tests [5, 12, 13, 14]. Three genes, chalcone synthase (CHS), pathogenesis-related protein (PRB), and polygalacturonase-inhibiting protein (PGIP), all involved in plant immune system, mapped in proximity to WM2.2 [12, 15, 16]. Molecular markers developed for this QTL are currently used for marker assisted selection (MAS) in common bean breeding programs [13, 17, 18].

Given the importance of this major QTL in WM resistance, narrowing the QTL interval and identifying candidate genes will provide a basic understanding of the genetic mechanisms of resistance to this disease, and enable the development of tightly linked markers to assist breeding and improvement of white mold resistance in common bean.

Methods

Plant materials and phenotyping

Two populations, R31, consisting of 105 F_{5:7} RILs, and Z0726–9, consisting of 86 F_{5:7} RILs, were used for a WM2.2 QTL-based sequencing study (Fig.S1). R31 was generated from a cross between Raven (seeds provided by Dr. James Kelly–Breeder at Michigan State University) and I9365–31, two black bean genotypes. Raven is known to be a highly susceptible cultivar to WM disease [19], while I9365–31 is a partially resistant germplasm line derived from an interspecific cross between *P. vulgaris* and *P. coccineus* [20]. The resistant genetic factors within I9365–31 are presumably from the *P. coccineus* resistant parent since it possesses high levels of resistance to WM and has multiple WM meta-QTL [9, 14, 20, 21]. Beside WM2.2, QTL on Pv05, Pv06, and Pv07 were also reported in R31 population [9].

The Z0726–9 population was derived from a cross between AN–37 and PS02–029C–20. AN–37 a RIL selected from the Aztec/ND88 RIL population, was released as the WM resistant pinto bean USPT-WM–1 [22]. AN37 is the putative source of the WM2.2 QTL and possesses WM3.1 resistance QTL as well [16]. The PS02–029C–20 (Matterhorn*4/NY6020–4) parent was reported to possess the WM8.3 QTL but not the WM2.2 QTL [11].

The field reaction to WM disease was evaluated at physiological maturity as described by Miklas et al. [7], based on a 1 to 9 scale where 1 is no disease development, and 9 is completely dead plants. The phenotypic data for the Z0726–9 population was collected during the 2009 and 2010 growing seasons. The mean 2009 field data was obtained from two ratings a few weeks apart with three replications, and the mean of 2010 field data was scored only once across three replications. The phenotypic data for the R31 population was initially published in Soule et al. [14]. The data was collected during 2000 and 2001 growing seasons, with three replications. The score for the individual plant within the plot with the most disease was recorded as well. In addition, the field score used for the QTL analysis was adjusted using plant stand as a covariate.

Classic QTL mapping

To confirm the existence of the WM2.2 QTL in both populations, a linkage map was re-generated using 5,400 SNPs [9] and additional Insertion-Deletion (InDel) markers developed within the physical boundaries of meta-QTL WM2.2. The InDels were designed through assembly and alignment of the sequence data from two Durango race genotypes, USPT-WM–1 (WM2.2 resistance, [22]) and Matterhorn (WM2.2 susceptible, [23]), against the *P. vulgaris* V2.1 reference genome sequence ([https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias = Org_Pvulgaris](https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Pvulgaris)). By comparing the mapped sequences of the two genotypes in IGV 2.3 [24], a total of 48 InDel markers were generated within the 3Mbp to 23Mbp interval on Pv02 with a minimum size of 6 bp and a mean of 19.7 bp. For population R31, all SNPs and InDels were used to generate a linkage map, but for population Z0726–9, only InDel markers were employed. Mapdisto v2 [25] was used to generate linkage groups based on a minimum logarithm of odds (LOD) score of 3.0 and maximum recombination frequency of 50%. ICMapping v4.1 [26] was used for an Inclusive Composite Interval Mapping (ICIM) QTL analysis [27]. The LOD value was calculated based on 1000 permutation tests at two experiment-wide error rates; $\alpha = 0.01$ and $\alpha = 0.05$ [28]. The significant QTL boundary was based on genetic intervals that passed the threshold of LOD score (3.16 for R31 and 2.01 for Z0726–9). Markers within QTL were further used to identify resistant and susceptible genotypes for bulked DNA sequencing.

Selecting High and Low bulks for DNA Isolation and sequencing

Genotypes with extreme WM resistance or susceptibility phenotypes, and possessing alleles associated with those phenotypes based on the QTL analysis, were selected for bulking. DNAs of the selected extreme lines from R31 and Z0726–9, along with their resistant parental lines, were extracted using MagMax magnetic bead purification protocol (<https://www.thermofisher.com>). The resistant and susceptible DNA bulks, along with the resistant parental lines, were sequenced (2x150bp paired-end reads on Illumina X10) at HudsonAlpha Institute for Biotechnology (<https://hudsonalpha.org/gsc/>).

Data processing and SNP calling

For each population, poor-quality bases were trimmed using Sickle [29]. Paired end reads were aligned against the resistant parent sequence using BWA-MEM [30]. The resistant parental reference genome was developed from

P. vulgaris reference genome, G19833 V2 using “Fasta Alternate Reference Maker” algorithms in GATK v3.3 [31]. After mapping, the Picard MarkDuplicates was used to remove duplicates for each of the bulks. Sequences were indexed using Samtools [32]. SNP variants for the susceptible bulk for each population were called using GATK Unified Genotyper v3.3 [31] with a minimum confidence threshold of 30 and filtered for a minimum read depth of five.

Variant identification

For each pool, we calculated the reference allele frequency as the ratio of reads that matched the reference allele to the total number of reads at that particular position using in-house scripts (<https://github.com/sujanmamidi/Pop-Seq>). The SNPs were categorized into fixed (the reference allele frequency is 100% in one bulk and 0% in the other bulk), shared (polymorphic in both bulks, the reference allele frequency is between 0 and 100), unique-resistant (polymorphic in only the resistant pool, but one allele is fixed in susceptible bulk), or unique-susceptible (polymorphic in only the susceptible pool, but one allele is fixed in resistant bulk).

The count of fixed SNPs for 10kb window with a 2kb slide (10kb–2kb) was performed. The significant regions of the genome were estimated at 95% and 99% percentile tail using 10,000 bootstraps. Candidate genes were selected within these physical boundaries of the genome. In addition, the potential effect of fixed SNPs was obtained using snpEFF [33].

Results

Phenotype

The mean WM score for population R31 was 5.6 with a range of 3.7 to 7.6, and for Z0726–9, the mean was 4.6 with a range of 2.2 to 7.3 (Fig. 1). Phenotypic values for both populations were normally distributed (KS test P-value at $\alpha_{0.01} = 1.63$), and the WM score mean between the resistant (score <5 for R31 and <4 for Z0726–9) and susceptible (score >6 for R31 and >5 for Z0726–9) haplotypes were significant (Table 1). Z0726–9 lines with resistant haplotypes showed significantly lower (more resistant) WM scores compared with R31 lines, which likely results from the presence of an additional QTL (WM8.3) contributed by the parent PS02–029C–20 [11].

QTL Mapping

Classic QTL analysis confirmed the existence of WM2.2 in both populations (Fig. 2 and 3). For R31, WM2.2 was detected within an interval of 37.5 to 39.5 cM, corresponding to a physical interval of 3.57–4.54 Mbp, while in Z0726–9 the QTL was detected within an interval of 11.5 to 20.5 cM corresponding to 10.40 to 43.24 Mbp. Maximum LOD score and additive effects were 12.65 and 0.70 for R31, and 6.63 and 0.60 for Z0726–9, respectively (Table S 1).

For R31, the SNP alleles from five markers including: ss715647659, 3.58 Mbp; ss715648336, 4.28 Mbp; ss715648335, 4.33 Mbp; ss715639861, 4.48 Mbp; and ss715648338, 4.54 Mbp, were used to select genotypes for resistant vs susceptible bulks (Table S2). For Z0726–9, seven InDel markers: WM2.2–02–10403733 (bp),

WM2.2-02-10425994, WM2.2-02-10436002, WM2.2-02-12724443, WM2.2-02-13107477, WM2.2-02-14728780 and WM2.2-02-16830328 were used to define genotypes for bulking.

Selection of bulks

Based on both phenotypic data and WM 2.2 marker alleles, two extreme bulks for each population were generated. DNA of 33 genotypes (R31: resistant $n = 8$, susceptible $n = 9$; Z0726-9: resistant $n = 8$, susceptible $n = 6$; plus two resistance parents) were extracted and pooled accordingly (Table 2). The mean of resistant and susceptible bulks for R31 were respectively 4.0 (with a range of 3.7 to 4.3) and 7.1 (with a range of 6.8 to 7.4). The mean values of the two bulks are significantly different (p -value = 9.14×10^{-31}). For Z0726-9, the mean of resistant and susceptible bulks were respectively 2.4 (with a range of 1.9 to 2.8) and 6.4 (with a range of 6.1 to 6.7). The mean values of the two bulks are significantly different (p -value = 2.11×10^{-17}).

Sequencing and variant detection

For the R31 population, 223 million reads ($\sim 55x$) were obtained from the resistant parent I9365-31, and 95.71% of these were mapped. For this population, 257 million reads for susceptible and 240 million reads for resistant bulks, were generated (Table 3). For the Z0726-9 population, 180 million reads ($\sim 45x$) were obtained for the resistant AN-37 parent and 95.67% of these were mapped. A total of 239 million reads for susceptible and 204 million reads for resistant bulks were obtained for this population. The resistant parent from each population served as reference-based parental genome sequence. Therefore, reads for each bulk were mapped to the appropriate parental genome sequence. After filtering, we identified $\sim 197k$ SNPs for the R31 bulks and $\sim 278k$ SNPs for the Z0726-9 bulks.

A total of 2,414 fixed SNP sites were detected between the resistant and susceptible pools for the R31 population. Two QTL based on significant windows (at 99 percentile) were identified in the euchromatic regions at Pv02:3.54-3.56 Mbp (with an average of 68 SNP per 10k window) and Pv02:4.27-4.56 Mbp (with an average 82 SNP per 10k window). The fixed SNP peak was located at 3.54 and 4.48 Mbp respectively (Fig. 4a). For the Z0726-9 population, there were 66,189 fixed SNPs between the pools. For this population, nine QTL based on significant windows (at 99 percentile) were identified in the heterochromatic region of Pv02 at 12.19 to 26.41 Mbp. Fixed SNP peaks were located at 12.20, 12.99, 13.58, 17.45, 19.14, 19.93, 21, 23.38 and 26.32 Mbp respectively in each window. The highest number of SNPs (233) was located at 13.83-13.92 Mbp significant window (Fig. 4b).

Candidate genes

For the WM2.2 region defined in the R31 population, five gene models were identified in the Pv02:3.54-3.56 Mbp interval, and 16 gene models were discovered in Pv02:4.27-4.56 Mbp interval (Table 4). Of these 21 genes, three gene models (*Phvul.002G044500*, *Phvul.002G046300*, and *Phvul.002G049000*) encoded a pentatricopeptide repeat (PPR) superfamily protein, a leucine-rich repeat (LRR) family protein, and an ortholog of gibberellin 2-oxidase 8, respectively. SNPs within each of these gene models have a potential modifier effect on WM2.2 resistance. For the Z0726-9 population, which has a much broader interval (12.19-26.41 Mbp), 12 gene models

were identified as candidates (Table 4). Clusters of LRR family proteins and pathogenesis-related thaumatin superfamily proteins in the Z0726–9 population were described to have a modifier effect. A domain analysis of the three LRR proteins in this interval revealed that only gene model *Phvul.002G079200*, an ortholog of TIR-NBS-LRR class of disease resistance proteins, has all the required NB-ARC domain amino acid signatures (GKTTL, GLPL, and MHD) necessary for the activation of a functional disease resistance gene [34].

Non-synonymous fixed sites

The scan of non-synonymous fixed sites using snpEff identified additional regions associated with WM disease in each population. For R31, gene models *Phvul.002G046600*, *Phvul.002G048200* and *Phvul.002G048900* [respectively orthologs of heat shock protein (HSP70), protodermal factor (PDF) and tyrosyl-DNA phosphodiesterase-related (TDP)] and for Z0726–9, gene models *Phvul.002G100600*, *Phvul.002G114500* and *Phvul.002G124300* [respectively orthologs of proxisom (PEX), squamosa promoter binding protein(SPL), and EF-TU receptor (EFR)] were annotated as non-synonymous fixed sites with moderate effects on WM. Among these gene models associated with missense variants, the role of HSP70 and EFR in plant immune response have been studied extensively.

Discussion

Advances in high throughput sequencing technologies enable the mapping of genetic factors associated with multiple traits and the subsequent identification of candidate genes [35, 36]. Next generation sequencing, NGS, is widely utilized for QTL and GWAS (genome wide association studies) analyses to identify the regions responsible for a specific phenotype. Both QTL and GWAS methods have advantages and disadvantages. Biparental QTL analysis can identify a strong genetic factor but the confidence interval is affected by the number of recombination events in the population. GWAS utilizes a high number of recombination events within the population leading to the identification of narrower QTL intervals compared to biparental QTL mapping. However, association mapping studies are constrained because many QTL may be observed that explain only a small portion of the phenotypic variation. To overcome the limitations of QTL mapping and association mapping, and to narrow the interval of an important QTL for WM resistance in common bean, we used bulked-segregant analysis (BSA) coupled with marker selection [11]. To optimize the BSA, we used marker data from a preliminary biparental QTL analysis along with phenotypic data to select lines with resistant or susceptible haplotypic state to select lines for the DNA bulks. Two RIL populations which possess WM resistance from different genetic backgrounds were evaluated. Population R31 was developed from a cross of two black bean parents, while population Z0726–9 was derived from crossing pinto and great northern parents [9, 11]. QTL WM2.2 was previously discovered in the genotypes, I9365–31 and USPT-WM–1, the parents of the R31 and Z0726–9 populations respectively. Given the WM2.2 intervals were large in these populations, it was of interest to fine map this QTL.

In this study, the R31 QTL was detected within the same PV02 region that meta QTL WM2.2 was previously mapped [9] but with a narrower 3.57–4.54 Mbp interval. For Z0726–9, the QTL was identified in a broad region that overlaps the heterochromatin region of the genome. This broad region also overlaps QTL in the BV and R31 populations [9]. This large region may have resulted from the low level of recombination in the Pv02 heterochromatic region that is estimated to extend from 8.0–27.1 Mbp in the reference genome of common bean

[36]. Fixed SNPs provide the greatest genotypic contrast between the two bulks. The significant peak interval in R31 was narrower than the Z0726–9 peak. This occurred because the R31 peak is located in the highly recombinogenic euchromatic region of Pv02 whereas the Z0726–9 peak is located in the low recombination heterochromatic region. This has been seen previously in common bean where the WM8.3 QTL was identified in the heterochromatic region of Pv08 across 10.75 Mb, while QTL WM7.1 mapped in the euchromatic region of Pv07 to a narrow 1.25 Mbp interval [11]. In one other comprehensive study of 771 QTL related to 161 unique sorghum traits, the heterochromatic region had a mean QTL density of 11QTL/0.5 Mbp compared to 7.5 QTL/0.5 Mbp in the euchromatic region [37]. Although QTL detected in the recombination-poor regions might complicate successful gene cloning, they can be very useful in MAS breeding programs. Moreover, the availability of a *P. vulgaris* reference genome sequence will help overcome the complication of gene cloning in these regions.

Combining the approaches of QTL with a greater number of markers and sequence-based introgression, it was determined that the meta-QTL WM2.2 consists of two independent QTL regions WM2.2a (= 1.67 Mbp) and WM2.2b (= 14.22 Mbp). WM2.2a, physically mapped in the euchromatic region, and WM2.2b in the heterochromatic region, may provide different WM tolerance/resistance mechanisms.

Using only significant fixed polymorphism sites between the two extreme groups of phenotypes allowed us to select candidate genes that were highly polymorphic between susceptible and resistance parent. Furthermore, using shared sites that differed by allele frequency of 0.3 or more, a minor QTL was detected within a small interval that corresponded with meta-QTL WM5.4 in the R31 population. This suggests that WM2.2 and WM5.4 may interact with each other in the R31 population to provide a higher level of resistance [9]. Previously in the fine mapping of the WM7.1 QTL, the resistant allele for QTL WM9.1 was also detected, and it was hypothesized that the interaction of the two loci provided a higher level of resistance to WM disease [11].

Our fine mapping revealed potential candidate genes associated with synonymous and/or non-synonymous fixed sites in WM2.2a and WM2.2b intervals. Pentatricopeptide repeat (PPR) proteins and gibberellin 2-oxidase are two important disease tolerance/resistance associated gene models detected in WM2.2a. PPR proteins share sequence features with NLR disease resistance genes particularly for their involvement in diversifying selection process. It has been shown that the birth and death process described for immunoglobulin genes [38], and later for plant R genes [39] is also the route for tandem repeat duplications in the PPR protein gene family [40]. Many studies have shown that PPR proteins contain many sequence-specific RNA binding sites involved in several RNA processing activities [41, 42, 43, 44, 45]. For example, the mRNA encoding PPR proteins are the target for miR400 that down regulates the PPR genes by cleaving their mRNAs and subsequently the plant becomes more susceptible to pathogenetic bacteria and fungi [46]. Moreover, recent studies on sequenced genomes of several plants revealed that the nucleotide binding site of PPR protein genes are similar to the NB-ARC nucleotide binding site of NLR genes which are part defense mechanism used by plants to respond to pathogen infection [47].

Gibberellin 2-oxidase is candidate gene that may explain WM disease avoidance mechanism associated with plant architecture. Gibberellins controls inter-node length and therefore indirectly the shape of the bean plant canopy. Shorter internode plants have an upright canopy that allows airflow through the field which in turns reduces the high humidity environment favored by the white mold pathogen. Long internode plants become prostrate on the ground which leads to a low, dense canopy with high humidity conditions favored by the pathogen. This leads to an increased severity of WM disease in the field [5, 48]. Gibberellin 2-oxidase inactivates gibberellins and effectively changes the status of this important hormone for internode extension. This feature

makes gibberellin 2-oxidase an important candidate gene because WM2.2 is associated with canopy porosity, canopy height, and lodging in the R31 population. Vasconcellos et al. [9], detected a QTL for canopy porosity, co-located (5.8 Mb) with WM2.2 in R31 population. Furthermore, WM2.2 was not detected in the greenhouse straw test used to detect physiological resistance. This further supports avoidance related genes as primary candidates. Moreover, significant fixed SNP sites detected in the WM2.2a interval were located inside the start site of the gibberellin 2-oxidase gene which makes it a reasonable candidate gene for disease avoidance traits.

Hsp70, a candidate gene in the WM2.2a interval, interacts with other plant defense genes [49]. HSPs are molecular chaperone proteins which play an important role in modulating the structure of disease resistance proteins, and are found to modulate Arabidopsis defense against pathogens [50]. In another study, Jelenska et al. [51], reported that the *P.syringae* pathogen uses the HopI1 effector to hijack Hsp70-associated chaperone activity. Under heat stress conditions, if excess HSP70 is provided, this effector is dispensable for *P.syringae* pathogenicity. This suggests searching for candidate genes associated with disease resistance, in addition to looking for NB-ARC genes in significant intervals, other genes which might be the target of a pathogen effector should also be considered. These genes/proteins might or might not have a R gene domain structure, but rather may be critical cellular proteins that directly or indirectly influence R-gene functions by modulating their structure or/and stability following their interaction with an effector [52].

The significant WM2.2b interval discovered in the Z0726–9 population was identified in the heterochromatic region. A cluster of genes encoding LRR proteins, typical of plant disease associated resistance genes, were detected in this region. This finding suggests a role in physiological resistance for WM2.2 in the Z0726–9 population. Partial physiological resistance is important when the severity of the disease overcomes the avoidance mechanisms provided by upright architecture or, reduced lodging. However, when all of the LRR clusters in this region (19.78–20.21 Mbp) were analyzed in Pfam [53], only a single TIR-NBS-LRR gene model (*Phvul.002G079200*) possessed all NB-ARC domain amino acid signatures necessary for a functional disease resistance protein. Among the gene models associated with non-missense variants in the WM2.2b interval, an EFR gene (*Phvul.002G124300*) is an important part of plant immune response. This receptor is known as a pattern-recognition receptor (PRR) with an extracellular domain harboring the LRR domain which recognizes and binds to the prokaryotic elongation factor EF-Tu [54]. This response prevents subsequent spread of the infection through the plant after the plant becomes infected by the pathogen. However, this receptor can only be effective if the pathogen expresses an excess level of EF-Tu [55].

Conclusions

Using additional markers for QTL analysis, and fixed sites for introgression, the WM 2.2 QTL was fine-mapped into two separate regions, WM2.2a (3.54–4.56 Mbp) and WM 2.2b (12.19 to 26.41 Mbp). QTL WM2.2a may trigger an avoidance mechanism, possibly through gibberellin 2-oxidase. When confronted with moderate to severe WM disease in the field, WM2.2b may induce physiological resistance via a traditional NLR reaction network. Classic QTL-linked indel markers described here can be used for marker assisted selection to combine avoidance and physiological resistance in common bean germplasm.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interest

The authors declare that they have no competing interest.

Funding

This publication was supported by the U.S. Department of Agriculture's (USDA) National Sclerotinia Initiative.

Authors' contributions

Atena Oladzad, Sujan Mamidi, Phillip McClean and Phillip Miklas designed and conceived the experiment. Phillip Miklas performed the experiments in the field and greenhouse, developed the populations, provided the phenotyping and linkage mapping data for the classic QTL analysis. Rian Lee designed and ran the InDel markers. Atena Oladzad, Sujan Mamidi, and Phillip McClean discussed the results and interpretation of the final data and Phillip Miklas provided suggestions to improve it.

Availability of data

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgements

Not applicable

References

1. Schwartz HF, Singh SP. Breeding common bean for resistance to white mold: A review. *Crop Science*. 2013;53:1832–1844.
2. Singh SP, Schwartz HF, Viteri D, Terán H, Otto K. Introgressing white mold resistance from *Phaseolus coccineus* PI 439534 to common pinto bean. *Crop Science*. 2014;54:1026–1032.
3. Kabbage M, Yarden O, Dickman MB. Pathogenic attributes of *Sclerotinia sclerotiorum*: switching from a biotrophic to necrotrophic lifestyle. *Plant Science*. 2015;23:53–60.

4. Schwartz HF, Steadman J. White mold. Bean production problems in the tropics. 2nd ed. CIAT Cali Colombia. 1989;211–230.
5. Miklas PN, Porter LD, Kelly JD, Myers JR. Characterization of white mold disease avoidance in common bean. *European Journal of Plant Pathology*. 2013;135:525–543.
6. Miklas P, Grafton K, Nelson B. Screening for partial physiological resistance to white mold in dry bean using excised stems. *Journal of the American Society for Horticultural Science*. 1992;117:321–327.
7. Miklas PN, Johnson WC, Delorme R, Gepts P. QTL conditioning physiological resistance and avoidance to white mold in dry bean. *Crop Science*. 2001;41:309–315.
8. Singh SP, Schwartz HF. Breeding common bean for resistance to diseases: a review. *Crop Science*. 2010;50:2199–2223.
9. Vasconcellos RC, Oraguzie OB, Soler A, Arkwazee H, Myers JR, Ferreira JJ, et al. Meta-QTL for resistance to white mold in common bean. *PLoS one*. 2017;12: e0171685.
10. Balasubramanian P, Conner R, McLaren D, Chatterton S, Hou A. Partial resistance to white mould in dry bean. *Canadian journal of plant science*. 2014;94:683–691.
11. Mamidi S, Miklas PN, Trapp J, Felicetti E, Grimwood J, Schmutz J, et al. Sequence-based introgression mapping identifies candidate white mold tolerance genes in common bean. *The plant genome*. 2016; 9:2.
12. Kolkman J, Kelly J. QTL conferring resistance and avoidance to white mold in common bean. *Crop Science*. 2003;43:539–548.
13. Miklas PN. Marker-assisted backcrossing QTL for partial resistance to *Sclerotinia* white mold in dry bean. *Crop Science*. 2007;47:935–942.
14. Soule M, Porter L, Medina J, Santana GP, Blair MW, Miklas PN. Comparative QTL Map for White Mold Resistance in Common Bean and Characterization of Partial Resistance in Dry Bean Lines VA19 and I9365–3 1. *Crop Science*. 2011;51:123–139.
15. Ender M, Kelly JD. Identification of QTL associated with white mold resistance in common bean. *Crop Science*. 2005;45:2482–2490.
16. Miklas PN, Larsen KM, Terpstra KA, Hauf DC, Grafton KF, Kelly JD. QTL analysis of ICA Bunsu-derived resistance to white mold in a pinto × navy bean cross. *Crop science*. 2007;47:174–179.
17. Ender M, Terpstra K, Kelly JD. Marker-assisted selection for white mold resistance in common bean. *Molecular Breeding*. 2008;21:149–157.
18. Porto ACM, Souza-Lopes F, Vasconcellos RC, Leite ME, Dos Santos JB. SSR markers linked to meta-qtls for white mold resistance in common bean. *Annual report of the bean improvement cooperative*. 2017;60:35–36.
19. Kelly J, Hosfield G, Varner G, Uebersax M, Haley S, Taylor J. Registration of ‘Raven’ black bean. *Crop science*. 1994;34:1406–1407.
20. Miklas P, Grafton K, Kelly J, Schwartz H, Steadman J. Registration of four white mold resistant dry bean germplasm lines: I9365–3 I9365–5 I9365–31 and 92BG–7. *Crop science*. 1998;38:1728–1729.
21. Schwartz HF, Otto K, Terán H, Lema M, Singh SP. Inheritance of white mold resistance in *Phaseolus vulgaris* × *P. coccineus* crosses. *Plant Disease*. 2006;90:1167–1170.
22. Miklas PN, Kelly JD, Steadman JR, McCoy S. Registration of pinto bean germplasm line USPT-WM–12 with partial white mold resistance. *Journal of Plant Registrations*. 2014;8:183–186.

23. Kelly J, Hosfield G, Varner G, Uebersax M, Taylor J. Registration of 'Matterhorn' great northern bean. *Crop science*. 1999;39:589–590.
24. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature biotechnology*. 2011;29:24.
25. Heffelfinger C, Fragoso CA, Lorieux M. Constructing linkage maps in the genomics era with MapDisto 2.0. *Bioinformatics*. 2017;33:2224–2225.
26. Meng L, Li H, Zhang L, Wang J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*. 2015;3:269–283.
27. Zeng ZB. Precision mapping of quantitative trait loci. *Genetics*. 1994;136:1457–1468.
28. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics*. 1994;138:963–971.
29. Joshi N, 2015. FJ. Sickle: A sliding-window adaptive quality-based trimming tool for FastQ files 2011. (Version 1.20)[Software].
30. Li H. Aligning sequence reads clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997. 2013
31. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*. 2010;20:1297–1303.
32. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079.
33. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*. 2012;6:80–92.
34. Sukarta OC, Sloopweg EJ, Govere A. Structure-informed insights for NLR functioning in plant immunity. *Seminars in cell and developmental biology Elsevier*. 2016;134–149.
35. McClean PE, Bett KE, Stonehouse R, Lee R, Pflieger S, Moghaddam SM, et al. White seed color in common bean (*Phaseolus vulgaris*) results from convergent evolution in the P (pigment) gene. *New Phytologist*. 2018;219:1112–1123.
36. Oladzad A, Porch T, Rosas JC, Moghaddam SM, Beaver J, Beebe SE, et al. Single and Multi-trait GWAS Identify Genetic Factors Associated with Production Traits in Common Bean Under Abiotic Stress Environments. *G3: Genes Genomes Genetics*. 2019;9:1881–1892.
37. Mace E, Jordan D. Integrating sorghum whole genome sequence information with a compendium of sorghum QTL studies reveals uneven distribution of QTL and of gene-rich regions with significant implications for crop improvement. *Theoretical and Applied Genetics*. 2001;123:169.
38. Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences*. 1997;94:7799–7806.
39. Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome research*. 1998;8:1113–1130.
40. Geddy R, Brown GG. Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. *BMC genomics*. 2007;8:130.

41. Fujii S, Small I. The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytologist*. 2012;191:37–47.
42. Laluk K, AbuQamar S, Mengiste T. The Arabidopsis mitochondria-localized pentatricopeptide repeat protein PGN functions in defense against necrotrophic fungi and abiotic stress tolerance. *Plant Physiology*. 2011;156:2053–2068.
43. Saha D, Prasad A, Srinivasan R. Pentatricopeptide repeat proteins and their emerging roles in plants. *Plant Physiology and Biochemistry*. 2007;45:521–534.
44. Schaper E, Anisimova M. The evolution and function of protein tandem repeats in plants. *New Phytologist*. 2015;206:397–410.
45. Yagi Y, Nakamura T, Small I. The potential for manipulating RNA with pentatricopeptide repeat proteins. *The Plant Journal*. 2014;78:772–782.
46. Park YJ, Lee HJ, Kwak KJ, Lee K, Hong SW, Kang H. MicroRNA400-guided cleavage of pentatricopeptide repeat protein mRNAs renders Arabidopsis thaliana more susceptible to pathogenic bacteria and fungi. *Plant and Cell Physiology*. 2014;55:1660–1668.
47. Sekhwal M, Li P, Lam I, Wang X, Cloutier S, You F. Disease resistance gene analogs (RGAs) in plants. *International journal of molecular sciences*. 2015;16:19248–19290.
48. Tivoli B, Calonnec A, Richard B, Ney B, Andrivon D. Current knowledge on plant/canopy architectural traits that reduce the expression and development of epidemics. *European Journal of Plant Pathology*. 2013;135:471–478.
49. Lee JH, Yun HS, Kwon C. Molecular communications between plant heat shock responses and disease resistance. *Molecules and cells*. 2012;34:109–116.
50. Noël LD, Cagna G, Stuttmann J, Wirthmüller L, Betsuyaku S, Witte CP, et al. Interaction between SGT1 and cytosolic/nuclear HSC70 chaperones regulates Arabidopsis immune responses. *The Plant Cell*. 2007;19:4061–4076.
51. Jelenska J, Van Hal JA, Greenberg JT. Pseudomonas syringae hijacks plant stress chaperone machinery for virulence. *Proceedings of the National Academy of Sciences*. 2010;107:13177–13182.
52. Su J, Spears BJ, Kim SH, Gassmann W. Constant vigilance: plant functions guarded by resistance proteins. *The Plant Journal*. 2018;93:637–650.
53. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families' database. *Nucleic acids research* 2004;32:138–141.
54. Zhang W, Fraiture M, Kolb D, Loffelhardt B, Desaki Y, Boutrot FFG, et al. Arabidopsis receptor-like protein30 and receptor-like kinase suppressor of bir1–1/evershed Mediate Innate Immunity to Necrotrophic Fungi. *The Plant Cell*. 2013;25:4227–4241.
55. Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, et al. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*. 2006;125:749–760.

Tables

Table 1. phenotypic means and t-test values for two populations

Population		Number	W.M score mean
R31	Resistant WM2.2 haplotype	32	4.5
	All other lines	73	6.07
	t-test		6.91×10^{-27}
	Susceptible WM2.2 haplotype	36	6.63
	All other lines	69	5.05
	t-test		1.86×10^{-26}
	Resistant WM2.2 haplotype	32	4.5
	Susceptible WM2.2 haplotype	36	6.63
	t-test		9.14×10^{-31}
Z0726-9	Resistant WM2.2 haplotype	24	3.34
	All other lines	62	5.06
	t-test		2.11×10^{-17}
	Susceptible WM2.2 haplotype	31	5.68
	All other lines	55	3.9
	t-test		2.97×10^{-20}
	Resistant WM2.2 haplotype	24	3.34
	Susceptible WM2.2 haplotype	31	5.06
	t-test		2.11×10^{-17}

Table 2. Summary of WM score for selected genotypes for bulks

	R31		Z0726-9	
	Resistant Bulk	Susceptible Bulk	Resistant Bulk	Susceptible Bulk
# of samples	8	9	8	6
Mean	4.01	7.07	2.43	6.38
Range	3.74-4.32	6.77-7.42	1.9-2.8	6.1-6.7

Table 3. Detailed of paired end reads and SNPs in the bulks of each populations

	R31		Z0726-9	
	Susceptible	Resistance	Susceptible	Resistance
QC-passed reads	257,137,995	240,874,709	239,456,111	204,777,649
% mapped	97%	96%	96%	95%
Fixed SNPs	2,414		66,189	
Shared	37,651		50,022	
Unique Resistant	14,998		19,863	
Unique Susceptible	12,749		15,771	

Table 4. Candidate genes identified

Population	Chrom	Significant window (Mb)	Average window	Pick Fixed	# Genes	Candidate	Gene	At define
R31			SNP	Site				
	2	3.54-3.56	68	3.54	5	AT3G18110.1	EMB1270	Pentatricopeptide repeat (PPR) superfamily protein
	2	4.27-4.56	82.22	4.48	16	AT4G21200.1	ATGA2OX8,GA2OX8	gibberellin 2-oxidase 8
						AT3G43740.2		Leucine-rich repeat (LRR) family protein
Z0726-9	2	12.19-12.46	200	12.2	15	AT1G69970.2	CLE26	CLAVATA3/ESR-RELATED 26
						AT5G23400.1		Leucine-rich repeat (LRR) family protein
	2	12.99-13.16	210	12.99	7	AT2G04030.1	AtHsp90	Chaperone protein htpG family protein
	2	13.83-13.92	232.92	13.85	5			
	2	17.45-17.47	201	17.45	2	AT1G15000.1	scpl50	serine carboxypeptidase-like 50
	2	19.14-19.28	224.78	19.14	4	AT1G31260.1	ZIP10	zinc transporter 10 precursor
	2	19.78-20.21	194.53	19.93	7	AT5G18820.1	EMB3007	TCP-1/cpn60 chaperonin family protein
						AT5G17680.1		disease resistance protein (TIR-NBS-LRR class)
	2	20.86-22.14	210.88	21	34	AT5G08470.1	PEX1	peroxisome 1
	2	22.93-25.23	208.8	23.38	110	AT4G38660.1		Pathogenesis-related thaumatin superfamily protein
						AT4G36010.2		Pathogenesis-related thaumatin superfamily protein
						AT4G36180.1		Leucine-rich receptor-like protein kinase family protein
						AT5G66330.1		Leucine-rich repeat (LRR)

							family protein
					AT3G12500.1	ATHCHIB	basic chitinase
					AT3G12500.1	ATHCHIB	basic chitinase
					AT2G17220.1		Protein kinase superfamily protein
					AT2G17220.1		Protein kinase superfamily protein
					AT2G39360.1		Protein kinase superfamily protein
2	25.86- 26.412	195.06	26.32	36	AT2G35610.1	XEG113	xyloglucanase 113
					AT3G47570.1		Leucine-rich repeat protein kinase family protein
					AT3G47570.1		Leucine-rich repeat protein kinase family protein
					AT3G47090.1		Leucine-rich repeat protein kinase family protein
					AT1G31420.2	FEI1	Leucine-rich repeat protein kinase family protein

Figures

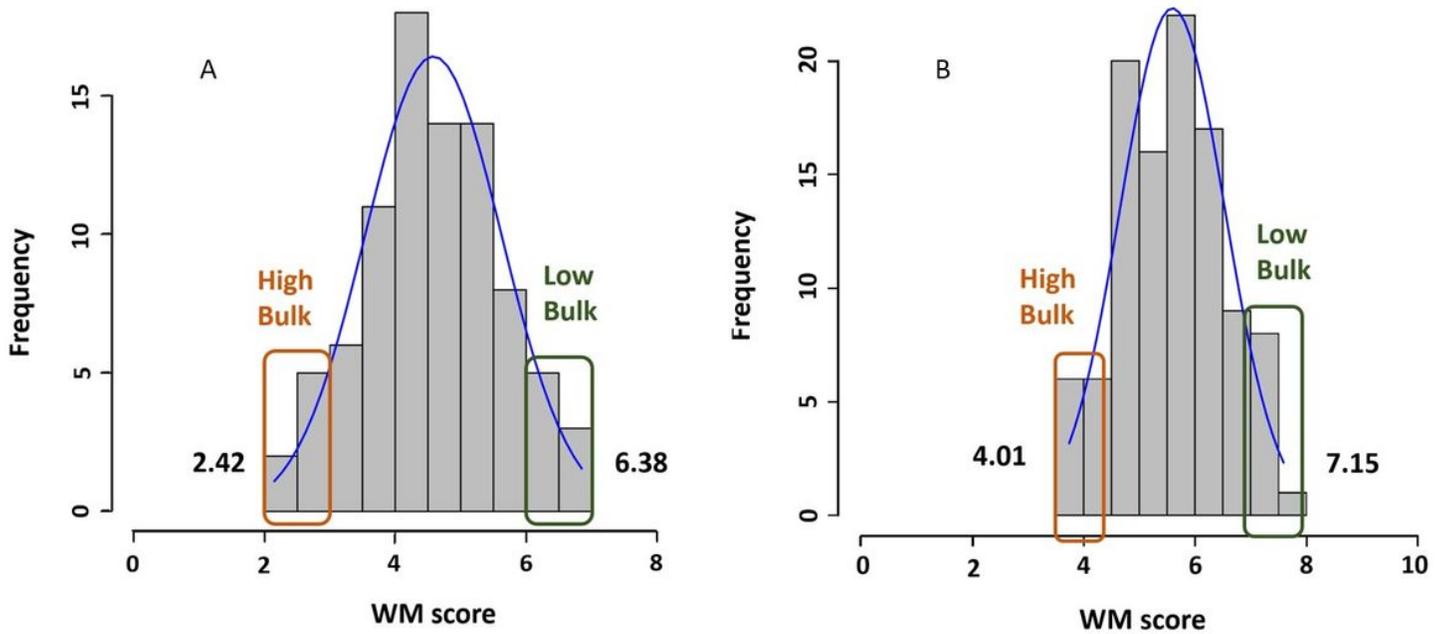


Figure 1

Histogram of WM phenotype score and selected progenies with highest and lowest score in Z0726-9 (a) and R31 population (b)

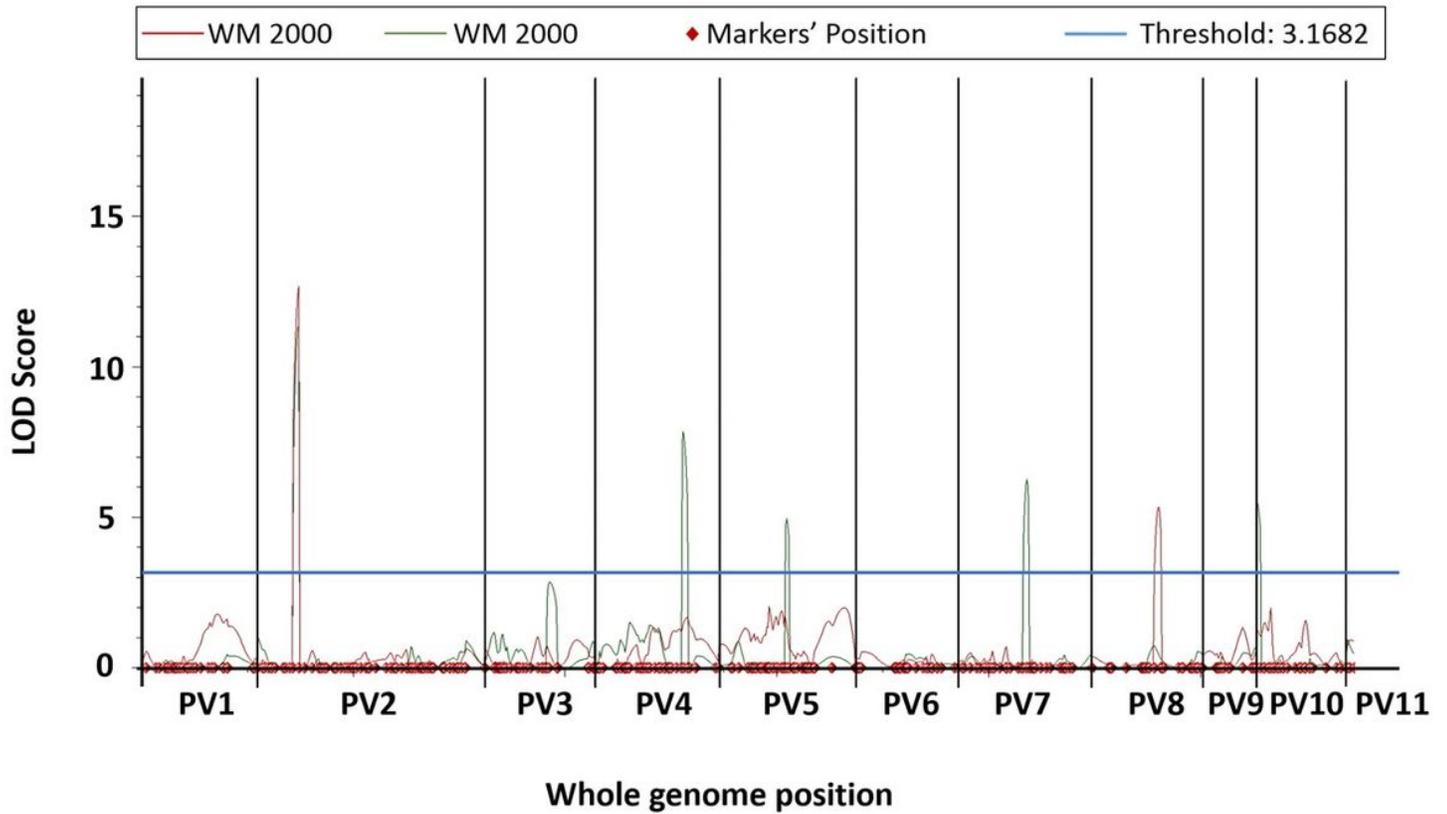


Figure 2

Classic QTL mapping of WM2.2 in R31 population

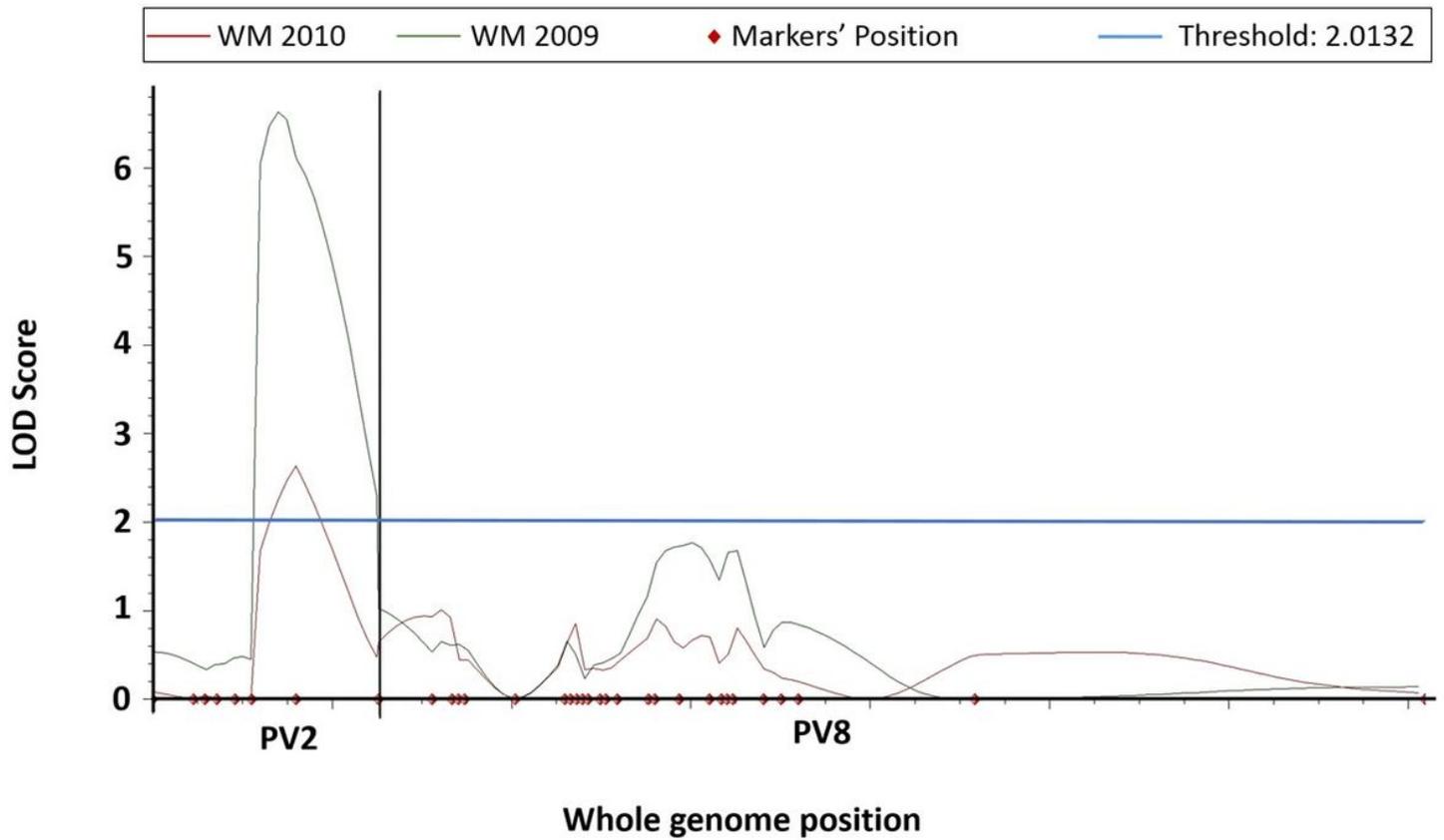


Figure 3

Classic QTL mapping of WM2.2 in Z0726-9 population

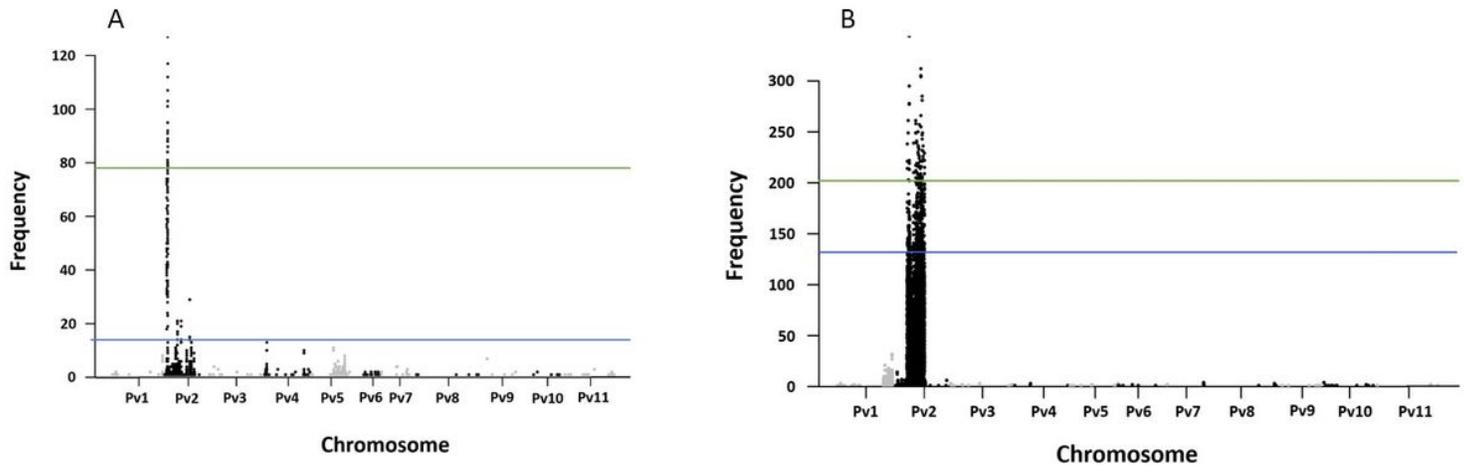


Figure 4

Fixed SNP counts in 10k-2k window (sliding-window). A) Population R31 B) Population Z0726-9

Supplementary Files

This is a list of supplementary files associated with this preprint. [Click to download.](#)

- [Supplementarytables.xlsx](#)
- [SupplemmentaryFigure.pptx](#)