

# Genome-wide Identification of Powdery Mildew Resistance in Common Bean

**Papias Hongera Binagwa**

Tuskegee University <https://orcid.org/0000-0001-6735-6325>

**Sy M. Traore**

Tuskegee University

**Marceline Egnin**

Tuskegee University

**Gregory C. Bernard**

Tuskegee University

**Inocent Ritte**

Tuskegee University

**Desmond Mortley**

Tuskegee University

**Kelvin Kamfwa**

University of Zambia

**Guohao He** (✉ [ghe@tuskegee.edu](mailto:ghe@tuskegee.edu))

Tuskegee University <https://orcid.org/0000-0001-5350-7786>

**Conrad C. Bonsi** (✉ [cbonsi@tuskegee.edu](mailto:cbonsi@tuskegee.edu))

Tuskegee University

---

## Research article

**Keywords:** genome, crop improvement, genotyping, loci, molecular breeding, GWAS

**Posted Date:** October 20th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-27731/v3>

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

[Read Full License](#)

---

# Abstract

**Background:** Genome-wide association studies (GWAS) have been utilized to detect genetic variations related to the powdery mildew (PM) resistance and several agronomic traits in common bean. However, its application in common bean and the PM interactions to identify genes and their location in the common bean genome has not been fully addressed.

**Results:** Genome-wide association studies (GWAS) through marker-trait association are useful molecular tools for the identification of disease resistance and other agronomic traits. SNP genotyping with a BeadChip containing 5398 SNPs was used to detect genetic variations related to resistance to PM disease in a panel of 211 genotypes grown under field conditions for two consecutive years. Significant SNPs identified on chromosomes Pv04 and Pv10 were repeatable, confirming the reliability of the phenotypic data scored from the genotypes grown in two locations within two years. A cluster of resistance genes was revealed on the Pv04 of common bean genome among which CNL and TNL like resistance genes were identified. Furthermore, two resistance genes *Phavu\_010G1320001g* and *Phavu\_010G136800g* were also identified on Pv10; further sequence analysis showed that these genes were homologs to the *Arabidopsis* disease resistance protein (RLM1A-like) and the putative disease resistance protein (*At4g11170.1*), respectively. Two LRR receptor-like kinases (RLK) were also identified on Pv11 in samples collected in 2018 only. Many genes encoding auxin-responsive protein, TIFY10A protein, growth-regulating factor 5-like, ubiquitin-like protein, cell wall protein RBR3-like protein related to PM resistance were identified nearby significant SNPs. These results suggested that the resistance to PM pathogen involves a network of many genes constitutively co-expressed and may generate several layers of defense barriers or inducible reactions.

**Conclusion:** Our results provide new insights into common bean and PM interactions, and revealed putative resistance genes as well as their location on common bean genome that could be used for marker-assisted selection, functional genomic study approaches to confirm the role of these putative genes; hence, developing common bean resistance lines to the PM disease.

## Background

Common bean (*Phaseolus vulgaris* L.) is a significant legume species among the pulse crops that play a major role in addressing global food security, environmental challenges and health diets [1]. Powdery mildew (PM) is one of the most ubiquitous plant diseases and it infects a variety of legumes including common bean. PM disease in common bean caused by pathogen *Erysiphe polygoni* DC accounts for extensive damage and significant yield losses 69% prior to flowering under the environment of warm temperatures (20-24°C), high humidity and shade [2,3]. Because it is an airborne disease, accurate identification and immediate taking actions are critical to effectively prevent the spread of PM thereby minimizing significant yield losses and quality of edible parts.

Initial symptoms appear as small and white talcum-like spots, which are most commonly seen on the upper surface of leaves [4]. As the symptoms develop, infected leaves gradually curl downward and change color from pale yellow to brown and ultimately abscise. Under severe conditions, the entire leaves and plants are covered by white cottony mycelia [5], which inhibit photosynthetic process through blocking light absorption by leaves [6] and decrease in the rate of photosynthetic carbon dioxide assimilation [7].

Several strategies are used to control PM disease, including application of fungicides, adjusting planting date to synchronize with periods of maximum sunlight exposure, and adoption of good cultural practices. However, these are expensive and not sustainable. For instance, fungicide treatments may not be effective in minimizing pathogen accumulation [8]. The development of resistant bean varieties is the most economical, efficient and ecological approach for the management of this disease-causing pathogen [9]. Screening sources of resistance and studying its inheritance in common bean genotypes have reported that resistant genotypes carry different resistance genes to the PM disease [4, 9-11].

Recently, susceptibility (S) genes have been used as an alternative source for the resistance to PM disease. The mildew resistance locus (mlo) is such S genes that promote the pathogen proliferation by suppressing the immune system; therefore, act as negative regulators of immunity [12,13]. Loss of function studies in some mlo genes have conferred a durable and broad-spectrum, recessively inherited resistance in Arabidopsis [14], cucumber [15], tomato [16,17] and grapevines [18]. Comparative genomics approach has revealed that five Mlo loci in the common bean genome were clustered in the clade V along with Arabidopsis orthologs underlying PM resistance [19]. However, the functionality of these Mlo loci has not been validated yet.

A Genome wide association study (GWAS) is a molecular tool used to identify specific genomic regions or loci governing simple to complex traits. GWAS can be used to determine if a genomic variant is associated with a trait of interest using either germplasm, segregation population, or a collection of diverse genotypes [20,21]. Various candidate genes or quantitative trait loci responsible for traits of interest have been identified using GWAS technology in different crops. For instance, in common bean several studies have reported identification of genomic regions and candidate genes associated with different traits for bruchid resistance [22], agronomic traits [23], drought tolerance [24], anthracnose, angular leaf spot and Fusarium wilt diseases [25,26] and symbiotic nitrogen fixation [27]. However, these studies have been conducted for different traits with different common bean cultivars in different geographical regions. In this research study, the aim was to reveal the association between genomic regions and the PM resistance using the GWAS approach in field grown cultivars. A collection of 211 common bean genotypes obtained from different resources and locations would provide enough recombination frequency for identification of causal loci related to PM resistance.

## Results

### Phenotyping for powdery mildew disease

Two trial sites were ideal experimental environments for testing PM disease resistant trait because of their higher disease pressure due to consecutively growth bean for more than five years resulting in the accumulation of *E. polygoni* inoculum season after season. The higher rainfall (>1000 mm) and higher relative humidity (>80%) also created conducive environment for disease development during these two bean growing seasons [28]. The environment was used to fully assess the PM resistance potentials of common bean genotypes in nature. As a result, highly significant ( $p < 0.001$ ) differences of resistant trait were observed among the 211 genotypes for powdery mildew in both 2017 and 2018 growing season [28]. Also, high significant interactions of genotype x environment and genotype x season were observed, indicating the resistance to PM disease is affected by environments (Table 1). Disease severity was higher in 2018 with maximum disease severity score of 9 compared to 2017 with a maximum disease severity score of 6 among 211 genotypes (Table 2). Disease severity of PM under natural infection was negatively correlated with yield at  $r = -0.24$  (data not shown), indicating that PM had a significant effect on yield of common bean.

Table 1. Analysis of variance in the collection of common bean genotypes for PM disease.

Source of variation	df	SS	MS	p-value
Replication	1	29.4603	29.4603	
Genotype (G)	210	668.3436	3.1826	<.001***
Environment (E)	1	624.8395	624.8395	<.001***
Season (S)	1	779.3892	779.3892	<.001***
GxE	210	337.5355	1.6073	<.001***
GxS	210	292.9858	1.3952	<.001***
ExS	1	105.0006	105.0006	<.001***
GxExS	210	224.8744	1.0708	0.093 ns
Residual	843	785.0397	0.9312	
Total	1687	3847.4686		

Note: df = degree of freedom; SS=Sum of Square; MS=Mean Square; \*\*\*=highly significant; ns=not significant. Significant level ( $p=0.05$ ).

Table 2. Means and variability of PM disease severity among 211 bean genotypes in 2017 and 2018.

Trait	Season	Min.	Max.	Mean of SE	P-value
Resistance to Powdery mildew (PM)	2017	1.00	6.00	1.51±0.02	<0.001
	2018	1.00	9.00	2.87±0.05	<0.001

## Population structure

Genetic relationship was revealed in the collection of 206 common bean accessions using STRUCTURE software. Based on Bayesian model, the genetic population structure was captured by describing the molecular variation in each subpopulation using a separate joint probability distribution over the observed sequence sites or loci. The model was used for association analysis to reduce false association due to unequal distribution of alleles among subpopulation. The model grouped 206 genotypes into three

clusters (K=3). The first cluster (K=1) consisted of 113 genotypes belonging to the Mesoamerican gene pool; the second cluster (K=2) consisted 72 genotypes belonging to the Andean gene pool and third cluster (K=3) consisted of 21 genotypes belonging to the admixture gene pool (Figure 1). Since this collection included some breeding lines derived from different sources of parents, there was a possibility that offspring combined genetic background from both Mesoamerican and Andean gene pools leading to these hybrids in the third cluster. The result suggested these subpopulations were associated with their genetic background. Interestingly, statistical analysis of disease severity scores based on genotypes in three structure groups showed the PM resistance was significantly different ( $p < 0.01$ ) between two gene pools, i.e. genotypes of Mesoamerican gene pool have significant higher resistances than those from Andean gene pool. However, genotypes in the admixture group have similar resistance with either gene pools depending on environment and season, suggesting the expression of resistant trait in admixture genotypes was easier influenced by environment.

Fig. 1. Population structure of 206 common bean genotypes with 5052 SNP markers. The color coding is as followed: red represent Mesoamerican gene pool (K=1), blue represent Andean gene pool (K=2), and green represent admixture of the two pools (K=3).

### Marker trait associations for powdery mildew disease

Genome-wide association analysis of the resistance to PM disease showed comparable results of significant SNPs between two the year's data. The most significant markers linked to PM resistance were located on Pv04 and Pv10 for both years, indicating the reliability of data set. A total of nine SNPs was detected as association with common bean powdery mildew resistance on the Pv04, Pv10, and Pv11 (Figure 2; Table 3). The SNP (ss715647913) at Pv10 was most significant ( $p = 7.22 \times 10^{-07}$ ) contributing the highest phenotypic variation of 12.28%. While the SNPs ss715650009 and ss715639212 had the highest significance levels representing 10.69% and 10.42% of phenotypic variation, respectively, on Pv04. There were two peaks of significant SNPs located in the range of 2.01-4.30 Mb and 45.01-45.02 Mb on Pv04, while only one significant SNP peak was detected from 40.57-41.13 Mb on Pv10. Significant SNPs were observed between 43.67-43.79 Mb on Pv11 only in 2018 (Table 3, Figure 2).

**Table 3.** Marker-trait association showing significant SNPs, chromosome position, and contribution to phenotypic variation for powdery mildew disease in 2017 and 2018.

SNP*	Chr.	Position (bp)	P-value	R <sup>2</sup> (%)
ss715647913	10	40,571,646	7.22 x 10 <sup>-07</sup>	12.28
ss715647918	10	40,624,302	9.36 x 10 <sup>-07</sup>	12.07
ss715645507	10	40,937,977	3.73 x 10 <sup>-06</sup>	10.93
ss715650009	4	2,013,385	5.22 x 10 <sup>-06</sup>	10.69
ss715648114	4	4,303,144	9.54 x 10 <sup>-05</sup>	8.29
ss715639212	4	45,014,435	1.11 x 10 <sup>-6</sup>	10.42
ss715648250	11	43,773,598	7.23 x 10 <sup>-06</sup>	11.01
ss715648263	11	43,673,759	2.29 x 10 <sup>-05</sup>	9.87
ss715648249	11	43,792,293	5.35 x 10 <sup>-05</sup>	10.81

SNP\* =Single Nucleotide Polymorphism code; Chr.=Chromosome; R<sup>2</sup> =phenotypic variation.

Fig. 2. A. Manhattan plots of two seasons for PM disease in 206 collection of bean genotypes using 5052 SNP markers. The horizontal line is the cut-off for significance and small green round indicated the significant markers. B. Distribution of SNPs in Pv04 and Pv10 and the range of SNP peaks.

### Identification of candidate genes associated with powdery mildew resistance

There were 181 coding genes observed within the interval of significantly associated SNPs on Pv04, 46 on Pv10, and 24 on Pv11. Using the BLAST analysis of these coding genes against protein database in GenBank, nine coding genes were considered as candidate resistance genes for the PM disease (Table 4). On the Pv04, three coding genes located at the telomeres were homologous to resistance genes *RPP13*, *TMV-N*, and *LRR receptor-likeserine/threonine protein kinase (LRR-RLK) RPK2*, while two coding genes located at the opposite telomere were homologous to one *LRR-RLK* and the transcription factor (TF) *MYB 87*. On Pv10, two coding genes were homologous with *RLM1A-like* and *At4g11170* putative resistance genes, while on Pv11 two other coding genes on Pv11 were similar with *LRR-RLK*. The sequences of these candidate resistance genes were listed in Additional file 1.

**Table 4.** Identified positional candidate genes linked to powdery mildew diseases resistance in dry bean.

Candidate gene	Position in Pv chromosome (bp)	Similarity	Gene accession	Gene Annotation
<i>Phavu_004G036200g</i>	Pv04:4014080 -4016937	95%	XM_017584784.1	<i>RPP13-like</i>
<i>Phavu_004G028900g</i>	Pv04:3126300 -3130090	99%	XM_017584233.1	<i>TMV N-like</i>
<i>Phavu_004G037500g</i>	Pv04:4141435 -4145345	99%	XM_014658246.2	<i>LRR-RLK (RPK2)</i>
<i>Phavu_004G173300g</i>	Pv04:45377489- 45382880	94%	XM_028069263.1	<i>LRR-RLK</i>
<i>Phavu_004G173500g</i>	Pv04:45394112- 45395393	100%	XM_014638647.2	<i>TF-MYB87</i>
<i>Phavu_010G1320001g</i>	Pv10:40237562- 40238278	95%	XM_022786275.1	<i>RLM1A-like</i>
<i>Phavu_010G136800g</i>	Pv10:40920056- 40929831	95%	XM_014660831.2	<i>At4g11170.1</i>
<i>Phavu_011G167800g</i>	Pv11:4379752 -43809211	98%	XM_028053325.1	<i>LRR-RLK</i>
<i>Phavu_011G169300g</i>	Pv11:4398238 -43999989	97%	XM_028054653.1	<i>LRR-RLK</i>

## Discussion

Most disease resistance in plant is a complex trait controlled by quantitative trait loci and influenced by environmental factors. The development of resistant varieties requires the identification of resistance genes as a prerequisite in plant breeding. Moreover, understanding the genetic basis of complex traits is needed for molecular breeding. GWAS is such a powerful approach for dissecting these complex traits and has been applied in many plant species including *Arabidopsis*, *Vigna unguiculata* and *V. radiata*. In this study, we used a nature population collected from different resources of common bean that possessed a large number of cross-over events. The BeadChip with 5398 SNPs provided enough required marker coverage on a diverse set of 206 accessions. The phenotype-genotype association was repeatable between the two years using GWAS approach, suggesting candidate gene identification pinned by significant associated SNPs was reliable.

Genome-wide association analysis of genes governing PM resistance resulted in the identification of nine candidate genes located on Pv04, Pv10 and Pv11 in the common bean genome that contains a total number of 28,134 coding genes ([https://plants.ensembl.org/Phaseolus\\_vulgaris/Location/Genome?](https://plants.ensembl.org/Phaseolus_vulgaris/Location/Genome?)

r=11:1-1000). Pv04 and Pv10 have the lesser coding genes (<2,000) compared to the rest of the chromosomes.

In this study, a coding gene (*Phavu\_004G036200g*) was identified at the telomere of Pv04 (between 4.014 - 4.016 Mb). BLAST analysis revealed that this gene is a homolog of the disease resistance gene RPP13. The RPP13 resistance gene encodes a coiled coil-nucleotide-binding site–leucine-rich repeat (CC-NBS-LRR, CNL) type of resistance protein and has been known to confer resistance to fungal diseases in plants, including resistance to downy mildew disease in *Arabidopsis* [29,30], resistance to powdery mildew in barley [31] and in wheat [32]. This RPP13-like gene located on Pv04 could play an important role in PM resistance in common bean. Recently, [4] have identified three loci related to PM resistance at the top arm of Pv04 in common bean based on the linkage mapping of three bi-parental segregation populations. Physically, these three loci were located between 0 to 1.09 Mb speculated from the cM of flanking DNA markers. Additional candidate gene (*Phavu\_004G001500*) related to PM resistance was identified by [11], which was located between 0.84-2.18 Mb on Pv04, suggesting the existence of a cluster of resistance genes in this region of Pv04. Indeed, in this study, BLAST analysis of coding genes around RPP13-like gene, detected several genes related to the PM resistance within 2 Mb interval between the two most significant SNPs, including *Phavu\_004G020000g* (senescence-associated carboxylesterase 101), *Phavu\_004G028900g* (TMV resistance protein N-like), and *Phavu\_004G037500g* (LRR-RLK, RPK2). Both TNL and CNL are involved in pathogen recognition but differ in signaling pathways. The TMV resistance protein N is encoded by Toll/interleukin-1 receptor-nucleotide-binding site–leucine-rich repeat type (TIR-NBS-LRR, TNL) gene [33]; activation of resistance response involving TNL requires several known general cofactors of disease resistance, including protein kinases [34]. At the opposite telomere, additional LRR-RLK and MYB87 could be involved in disease response since MYB87 functions as a regulator of genes affecting cell wall organization and remodeling [35]. It is probable that upon pathogen infection, the constitutive expression of this set of genes located on Pv04 could show the characterization of quantitative resistance trait to the PM disease in common bean. Further investigation is needed to elucidate the molecular mechanism of the cooperation between TNL and CNL pathways: whether TMV-like gene mediating the resistance to PM disease requires the protein kinase LRR-RLKs to induce the effector triggered immunity (ETI), or requires alternative spliced transcripts to promote resistance proteins that can specifically recognize the pathogen molecular elicitors [36].

Based on the two years' data, two other candidate resistance genes, *Phavu\_010G1320001g* and *Phavu\_010G136800g*, were identified at the bottom arm of Pv10 with 0.6 Mb apart nearby the significant SNPs. The former candidate gene was homologous with the disease resistance protein RLM1A-like, while the latter hit the putative disease resistance protein (*At4g11170.1*). The *RLM1A* gene confers resistance in *Arabidopsis* against *Leptosphaeria maculans*, a fungus pathogen, and is involved in the first layer of defense through the callose deposition acting as a temporary cell wall in response to pathogen attack [37-39]. The gene *At4g11170* refers to disease resistance gene *RMG1* (Resistance Methylated Gene 1) [40]. Expression of this gene is controlled by DNA methylation on its promoter region. The *RMG1* promoter region is constitutively demethylated by active DNA demethylation mediated by the DNA glycosylase ROS1 [40]. Both identified candidate genes encoded TNL resistance protein and their

physically closely located on Pv10. However, they did not show significant similarity in either DNA or amino acid sequences, suggesting they were not derived from the duplication event. These two clustered resistance genes may meet the digenic requirement for functional resistance as observed in the RPP2 that consists of a complex of TIR-NB-LRR genes for defense response [41].

Additional candidate resistance genes were detected on Pv11 based on the significant SNPs from the association analysis in 2018. Both genes were annotated as LRR receptor-like serine/threonine protein kinase (*At1g5430* and *At1g56130*). LRR receptor-like kinases (RLK) represent one of the largest protein families in plant [42]. The LRR-RLK proteins in plant are involved in the plant signaling pathway regulating pathogenic defense responses [43]. LRR domain of RLK has undergone an accelerated evolution that generated numerous cell surfaces and cytoplasmic receptors to interact with a diverse group of proteins for the specificity of pathogen recognition [43]. LRR receptor-like serine/threonine protein kinase were identified as candidate gene conferring the resistance to apple scab as demonstrated by [44]. In this study, LRR-RLKs were identified on Pv11 as associated with PM resistance based on one-year data. Although resistance loci were detected on Pv11 using different genotype materials of common bean [4,45], further study of the interaction between LRR-RLK and PM pathogen will help with understanding the mode of action of these resistance genes. Nevertheless, many genes encoding auxin-responsive protein, TIFY10A protein, growth-regulating factor 5-like, ubiquitin-like protein, cell wall protein RBR3-like protein were linked to significant SNPs, suggesting that the resistance to PM pathogen involves a network of many genes constitutively co-expressed and generates several layers of defense barriers or inducible reactions. The function of the candidate genes identified in this study needs to be elucidated by performing loss of function study.

## Conclusions

Common bean is affected by PM disease causing significant yield losses. In this study, we used the GWAS technology to identify SNPs associated with the resistance to the PM disease. The most significant SNPs were identified on Pv04 and Pv10. Further sequence analysis showed the presence of a clustered of resistance genes of the Pv04 of common bean among which several genes have been shown to be associated with the PM disease resistance. Moreover, two resistance genes were also identified on Pv10. Our results revealed putative resistance genes as well as their locations on the common bean genome that could be utilized for marker-assisted selection, functional genomic study approaches to confirm the role of these putative genes; therefore, provide the benefit in breeding program for common bean improvement.

## Methods

### Plant materials and field experiment

Seeds of 211 common bean accessions were originally sourced from different germplasm centers [27]. Among them, 184 accessions seeds were originally acquired from the common bean germplasm

repository at International Center for Tropical Agriculture (CIAT) ([www.ciat.cgiar.org](http://www.ciat.cgiar.org)) and the remaining 27 were derived from Tanzania Agricultural Research Institute (TARI) Selian center where they were originally collected from Ethiopia (12), Kenya (10), Tanzania (3), and Rwanda (2) ([www.tari.go.tz](http://www.tari.go.tz)). These genotypes represented two gene pools based on their origin centers, Mesoamerican origin in southern Mexico and Guatemala and Andean origin in Peru and Columbia [27]. The experimental materials were planted in northern Tanzania at medium altitudes of 107 m above sea level (a.s.l) of S03°21.690' and E36°37.879' (Selian Agricultural Research Institute abbreviated SARI) and at low altitudes of 992 m a.s.l of S03°19.905' and E037°14.067' (Tanzania Coffee Research Institute commonly called Lyamungo) in 2017 and 2018. In both seasons and agro-ecologies, experiments were laid-down in Incomplete Block Design with two replications for 211 genotypes. The plot size was 3.2 m long and 1.5 m wide with 4 rows per plot, spaced 0.5 m and 0.2 m under rainfed conditions. Data were collected from the middle two rows.

### **Phenotypic data analysis**

The PM disease severity scores were evaluated using 1-9 scale where 1 being non-pathogenic and 9 pathogenic following the scoring rubric of CIAT [46,47]. The evaluation of score was conducted twice i.e. at R-6 and R-8 developmental stages and the mean of two evaluations was used for downstream analyses. Analysis of Variance (ANOVA) was used to test the interactions of genotype x environment, genotype x season and genotype x environment x season. Protected Least Significant Differences (LSD) of ( $p=0.05$ ) were used for comparison of genotypes [48,49]. For each accession, mean of disease scores obtained from two replications and two environments in each year was used for the phenotype-genotype association study.

### **Genotyping data analysis**

Total of 15 seeds per genotype were shipped from Tanzania to Tuskegee University through USDA-Animal and Plant Health Inspection Services (APHIS) with permit number P587-180801-005 and phytosanitary certificate number 00310248 as per import/export plant material regulations. Seeds were planted in a pot with diameter of 15 cm in the greenhouse at the George Washington Carver Agricultural Experiment Station (GWCAES) of Tuskegee University Research Center, Alabama. Due to the failure of germination in five accessions, DNA samples were extracted from trifoliolate leaves of 206 accessions using Wizard® Genomic DNA Purification Kit from Promega Corporation (Madison, WI). The concentration of each DNA sample was determined using NanoDrop Spectrophotometer (NanoDrop 2000) and the quality of DNA samples was also assessed on 0.8% agarose gel. These 206 DNA samples were genotyped at Heflin Center Genomic Core Lab at the University of Alabama in Birmingham, AL, using the Next generation Sequencing (NGS) Method of Illumina BARCBean6K\_3 BeadChip with 5398 SNP markers [50]. All markers were distributed across 11 pairs of common bean chromosomes ( $2n=22$ ) and the Illumina BeadStation 500G was used to scan the BeadChips. The SNP calling was conducted with the genotyping modules of GenomeStudio v2.0.4 of 2018 available as a free download software.

### **Population structure analysis**

The Bayesian model-based clustering method was applied using STRUCTURE 2.3.4 software [51] to determine population in this collection. The admixture model with independent allele frequencies without prior population information was used for simulation. The STRUCTURE software was set at burn-in period length of 50,000 and after burn-in 50,000 Markov chain Monte Carlo (MMC) repetition were set five times. For joint inference of population substructure, the kinship (K) set at the range of 1 to 10 with five number of iterations runs for each kinship. The ideal number of sub-populations was determined using K method [52] implemented in the HARVESTER software [53].

### Marker-trait association analysis

Filtering the monomorphic SNP markers, and ones with minor allele frequency (MAF) <2% with 6.4% SNPs missing, 5052 SNP markers were retained for population structure analysis and association analyses with TASSEL 5.0 software. The following mixed linear model (MLM) was used:

$$Y = \gamma\alpha + \rho\beta + k\mu + \varepsilon$$

Where, Y is Phenotype of each genotype;  $\gamma$  is the fixed effects of the SNP;  $\rho$  is the fixed effect of population structure from principle component analysis results; k is the random effect of kinship relative;  $\varepsilon$  is the term error under normally distribution with mean =0 and variance  $\delta^2$ . The statistical model was used to test for trait-marker associations [54].

### BLAST analysis

Significant SNPs were identified using the manhattan plots for two bean growing season and the cut-off value 5 of  $-\log_{10}(\text{p-value})$  was used to identify the most significant SNPs associated with PM resistance. The coding genes within the interval of significantly associated SNPs were used as queries to search for putative homologous proteins in the EnsemblPlants release 46 version [55] and ([https://plants.ensembl.org/Phaseolus\\_vulgaris/Location/Genome?r=11:1-1000](https://plants.ensembl.org/Phaseolus_vulgaris/Location/Genome?r=11:1-1000)). The coding genes were annotated as candidate genes when they were homologous with genes, kinases, and transcription factors with nucleotide identity >90% and associated with disease resistance. The location of coding genes hitting the putative proteins was used as the position of candidate resistance genes on the corresponding chromosomes.

## Abbreviations

CC-NBS-LRR: coiled coil-nucleotide-binding site–leucine-rich repeat; TIR-NB-LRR: Toll/interleukin-1 receptor-nucleotide-binding site–leucine-rich repeat; ANOVA: Analysis of Variance; MLM: mixed linear model; ETI: effector triggered immunity; GWAS: genome-wide association studies; MAF: minor allele frequency; NGS: next generation sequence; PM: Powdery mildew; RLK: receptor like kinases; RMG: resistance methylated gene; RPP: resistance to *Peronospora parasitica*; SNP: single nucleotide polymorphism; TMV: tobacco mosaic virus;

# Declarations

## Acknowledgement

Authors are grateful to George Washington carver agricultural experiment station, Tanzania Agricultural Research Institute and Plant genomic lab Tuskegee University for greenhouse experimentation, phenotyping and genotyping respectively.

## Author' contributions

PHB, GH and CKB designed the research study and developed the manuscript. SMT and KK analyzed data, ME, GCB, IR and DM commented and corrected the manuscript. All authors proofread and approved the final version.

## Funding

This research work was funded by USDA-NIFA-Evans-Allen through Tuskegee University, Government of Tanzania and CIAT/PABRA through Tropical Legume III project.

## Availability of data and materials

Additional file 1: List of sequences of candidate resistance genes to PM in common bean.

## Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Competing interests

The authors declare that they have no competing interest

## Authors details

<sup>1</sup>Integrative Biosciences (IBS). Ph.D. Program, Tuskegee University, AL 36088 USA. <sup>2</sup>Department of Agricultural and Environmental Sciences, Tuskegee University, AL 36088 USA

<sup>3</sup>Department of Plant Sciences, School of Agricultural Sciences, University of Zambia, P.O Box 32379, Great East Road, Lusaka, Zambia.

# References

1. Calles T. The International year of pulses: What are they and why are they important? *Agriculture for Development*. 2016;27:40-42.
2. Rezende VF, Ramalho MAP, Corte HR. Genetic control of common bean (*Phaseolus vulgaris*) resistance to powdery mildew (*Erysiphe polygoni*). *Genet Mol Biol.* 1999;22(2):233–6. [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1415-47571999000200017&lng=en&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47571999000200017&lng=en&tlng=en)
3. Schwartz H, Stedman J, Hall R, Forster R. Compendium of bean diseases 2nd Edition. EDS. . 23:30-41. *Am Phytopathol Soc*. 2005;23:30–41.
4. Murube E, Campa A, Ferrera J. Identification of new resistance sources to powdery mildew and the genetic characterization of resistance in three common bean genotypes. *Crop pasture Sci*. 2017;68(11):1006–12.
5. Pernezney K, Stall W. Powdery mildew of vegetables. Gainesville; 2005.
6. Yamashita TT. Powdery mildew – symptoms, causes and treatment. The future of agriculture. 2019;<https://www.fusion360ag.com/powdery-mildew-symptoms-causes-treatment/>.
7. Magyarosy AC, Schumann P, Buchanan BB. Effect of powdery mildew infection on Photosynthesis by leaves and chloroplasts of sugar beets. *American Society of Plant Biologists*. 1976;DOI: <https://doi.org/10.1104/pp.57.4.486>.
8. Buruchara R, Mukankusi C, Ampofo K. Bean disease and pest identification and management. Handbooks for small-scale seed producers. Kampara: International Center for Tropica Agriculture; 2010;p67.
9. Trabanco N, Perez-Vega E, Campa A, Rubiales D, Ferreira J. Genetic resistance to powdery mildew in common beans. *Euphytica*. 2012;186:875–82.
10. Schwartz H, Katherman M, Thung M. Yield response and resistance of dry beans to powdery mildew in Columbia. *Plant Dis*. 1981;65:737–8.
11. Campa A, Ferreira JJ. Gene coding for an elongation factor is involved in resistance against powdery mildew in common bean. *Theor Appl Genet*. 2017;130:849-860.
12. Pavan S, Jacobsen E, Visser RGF, et al. Loss of susceptibility as a novel breeding strategy for durable and broad-spectrum resistance. *Mol Breeding*, 2010;25:1. <https://doi.org/10.1007/s11032-009-9323-6>.
13. Langner T, Kamoun S, and Belhaj, K. **CRISPR Crops: Plant Genome Editing Toward Disease Resistance**. *Annual Review of Phytopathology*, 2018;56:479-512. <https://doi.org/10.1146/annurev-phyto-080417-050158>.
14. Consonni C, Humphry ME, Hartmann HA, Livaja M, Durner J, Westphal L, et al. Conserved requirement for a plant host cell protein in powdery mildew pathogenesis. *Nat Genetics*, 2016;38:716-720.
15. Nie J, Wang Y, He H, Guo C, et al. Loss-of-function mutations in CsMLO1 confer durable powdery mildew resistance in cucumber (*Cucumis sativus* L.) *Front. Plant Sci*. 2015;<https://doi.org/10.3389/fpls.2015.01155>.

16. Andolfo, G., Iovieno, P., Ricciardi, L. et al. Evolutionary conservation of *MLO* gene promoter signatures. *BMC Plant Biol.* 2019;19:150. <https://doi.org/10.1186/s12870-019-1749-3>.
17. Zheng Z, Appiano M, Pavan S, Bracuto V, Ricciardi L, Visser RGF, et al., Genome-wide study of the tomato SIMLO gene family and its functional characterization in response to the powdery mildew fungus *Oidium neolyopersici*. *Front Plant Sci.* 2016;7:380.
18. Pessina S, Pavan S, Catalano D, Gallotta A, Visser RG, Bai Y, et al. Characterization of the *MLO* gene family in Rosaceae and gene expression analysis in *Malus domestica*. *BMC Genomics.* 2016;15:618. <https://doi.org/10.1186/1471-2164-15-618>.
19. Rispaill N, Rubiales D. Genome-wide identification and comparison of legume *MLO* gene family. *Sci Rep.* 2016;6:32673. <https://doi.org/10.1038/srep32673>.
20. Challa S, Neelapu N. Genome-Wide Association Studies (GWAS) for Abiotic Stress Tolerance in Plants. In: *Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants* Visakhapatnam. 2018;135–50.
21. Liu HJ, Yan J. Crop genome-wide association study: a harvest of biological relevance. *Plant J.* 2019;97(1):8–18.
22. Tigist SG, Melis R, Sibiya J, Amelework AB, Keneni G, Tegene A. Population Structure and Genome-Wide Association Analysis of Bruchid Resistance in Ethiopian Common Bean Genotypes. *Crop Sci.* 2019;59(4):1504–15. <https://onlinelibrary.wiley.com/doi/abs/10.2135/cropsci2018.09.0559>
23. Kamfwa K, Cichy KA, Kelly JD. Genome-Wide Association Study of Agronomic Traits in Common Bean. *Plant Genome.* *Plant Genome*, 2015;8(2):plantgenome2014.09.0059. <http://doi.wiley.com/10.3835/plantgenome2014.09.0059>
24. Hoyos-Villegas V, Song Q, Kelly JD. Genome-wide Association Analysis for Drought Tolerance and Associated Traits in Common Bean. *Plant Genome.* 2017;10(1):plantgenome2015.12.0122.
25. Perseguini JM, KC, Oblessuc PR, Rosa JRBF, Gomes KA, Chiorato AF, Carbonell SAM, et al. Genome-Wide Association Studies of Anthracnose and Angular Leaf Spot Resistance in Common Bean (*Phaseolus vulgaris* L.). Parida SK, editor. *PLoS One*, 2016;11(3):e0150506. <https://dx.plos.org/10.1371/journal.pone.0150506>
26. Leitão ST, Malosetti M, Song Q, Eeuwijk F, et al. Natural variation in Portuguese common bean germplasm reveals new sources of resistance against *Fusarium oxysporum* f. sp. *phaseoli* and resistance-associated candidate genes. *Phytopathology*, 2020; <https://doi.org/10.1094/PHTO-06-19-0207-R>.
27. Kamfwa K, Cichy KA, Kelly JD. Genome-wide association analysis of symbiotic nitrogen fixation in common bean. *Theor Appl Genet.* 2015;128(10):1999–2017.
28. Binagwa PH, He G, Egnin M, Bernard GC, Mortley D, Bonsi CK. Evaluating natural infection of fungal, bacterial and viral pathogens to dry bean genotypes under field conditions. *J. of Plant Breed. and Crop Sc.* 2020;12:70–90. <https://doi.org/10.5897/JPBCS2019.0866>.
29. Bittner-Eddy PD, Crute IR, Holub EB, Beynon JL. RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora*

- parasitica. *Plant J.* 2000;21(2):177–88. <http://doi.wiley.com/10.1046/j.1365-313x.2000.00664.x>
30. Rose L. The population genetics and functional analysis of the Pto disease resistance gene in *Lycopersicon* spp., and the RPP13 gene in *Arabidopsis*. University of California at Davis; 2002. <https://elibrary.ru/item.asp?id=5438682>
31. Wu Y, Ma X, Pan Z, Kale SD, Song Y, King H, et al. Comparative genome analyses reveal sequence features reflecting distinct modes of host-adaptation between dicot and monocot powdery mildew. *BMC Genomics.* 2018;19(1):705. <https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-018-5069-z>
32. Wu P, Hu J, Zou J, Qiu D, Qu Y, Li Y, et al. Fine mapping of the wheat powdery mildew resistance gene Pm52 using comparative genomics analysis and the Chinese Spring reference genomic sequence. *Theor Appl Genet.* 2019;132(5):1451–61. <http://link.springer.com/10.1007/s00122-019-03291-7>
33. Mestre P, Baulcombe DC. Elicitor-Mediated Oligomerization of the Tobacco N Disease Resistance Protein. *Plant Cell.* 2006;18(2):491–501. <http://www.plantcell.org/lookup/doi/10.1105/tpc.105.037234>
34. Jin H, Liu Y, Yang K-Y, Kim CY, Baker B, Zhang S. Function of a mitogen-activated protein kinase pathway in N gene-mediated resistance in tobacco. *Plant J.* 2003 ;33(4):719–31. <http://doi.wiley.com/10.1046/j.1365-313X.2003.01664.x>
35. Fujiwara S, Mitsuda N, Nakai Y, Kigoshi K, Suzuki K, Ohme-Takagi M. Chimeric repressor analysis identifies MYB87 as a possible regulator of morphogenesis via cell wall organization and remodeling in *Arabidopsis*. *Biotechnol Lett.* 2014;36:1049-1057.
36. Song S, Huang H, Gao H, Wang J, Wu D, Liu X, et al. Interaction between MYC2 and ETHYLENE INSENSITIVE3 Modulates Antagonism between Jasmonate and Ethylene Signaling in *Arabidopsis*. *Plant Cell.* 2014;26(1):263–79. <http://www.plantcell.org/lookup/doi/10.1105/tpc.113.120394>
37. Staal J, Kaliff M, Bohman S, Dixelius C. Transgressive segregation reveals two *Arabidopsis* TIR-NB-LRR resistance genes effective against *Leptosphaeria maculans*, causal agent of blackleg disease. *Plant J.* 2006;46(2):218–30. <http://doi.wiley.com/10.1111/j.1365-313X.2006.02688.x>
38. Persson S, Paredes A, Carroll A, Palsdottir H, Doblin M, Poindexter P, et al. Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. *Proc Natl Acad Sci.* 2007;104(39):15566–71. <http://www.pnas.org/cgi/doi/10.1073/pnas.0706592104>
39. Delgado-Cerezo M, Sánchez-Rodríguez C, Escudero V, Miedes E, Fernández PV, Jordá L, et al. *Arabidopsis* Heterotrimeric G-protein Regulates Cell Wall Defense and Resistance to Necrotrophic Fungi. *Mol Plant.* 2012;5(1):98–114. <https://linkinghub.elsevier.com/retrieve/pii/S167420521460650X>
40. Yu A, Lepere G, Jay F, Wang J, Bapaume L, Wang Y, et al. Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proc Natl Acad Sci.* 2013;110(6):2389–94. <http://www.pnas.org/cgi/doi/10.1073/pnas.1211757110>
41. Sinapidou E, Williams K, Nott L, Bahkt S, Tör M, Crute I, et al. Two TIR:NB:LRR genes are required to specify resistance to *Peronospora parasitica* isolate Cala2 in *Arabidopsis*. *Plant J.* 2004;38(6):898–

909. <http://doi.wiley.com/10.1111/j.1365-313X.2004.02099.x>
42. Gou X, He K, Yang H, Yuan T, Lin H, Clouse SD, et al. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in *Arabidopsis thaliana*. *BMC Genomics*. 2010;11(1):19. <http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-11-19>
43. Afzal AJ, Wood AJ, Lightfoot DA. Plant Receptor-Like Serine Threonine Kinases: Roles in Signaling and Plant Defense. *Mol Plant-Microbe Interact*. 2008;21(5):507–17. <http://apsjournals.apsnet.org/doi/10.1094/MPMI-21-5-0507>
44. Padmarasu S, Sargent DJ, Patocchi A, Troggio M, Baldi P, Linsmith G, et al. Identification of a leucine-rich repeat receptor-like serine/threonine-protein kinase as a candidate gene for Rvi12 (Vb)-based apple scab resistance. *Mol Breed*. 2018;38(6):73. <http://link.springer.com/10.1007/s11032-018-0825-y>
45. Pérez-Vega E, Trabanco N, Campa A, Ferreira JJ. Genetic mapping of two genes conferring resistance to powdery mildew in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet*. 2013;126(6):1503–12. <http://link.springer.com/10.1007/s00122-013-2068-y>
46. Van Schoonhoven A, Pastor-Corrales M. Standard system for the evaluation of bean germplasm. Cali; 1987.
47. CIAT-Kawanda. Training manual on phenotyping diseases of common bean (*Phaseolus vulgaris* L.) and characterization of the disease pathogens. 2013;Kampala:CIAT.
48. Clewer AG, Scarisbrick DH. Practical statistics and experimental design for plant and crop science. Wiley. Chichester, New York: Wiley; 2001.
49. Liu S, Yu K, Park S. Development of STS markers and QTL validation of common bacterial blight resistance in common bean. *Plant Breed*. 2008;127:62–8.
50. Song Q, Jia G, Hyten DL, Jenkins J, Hwang EY, Schroeder SG, et al. SNP assay development for linkage map construction, anchoring whole-genome sequence, and other genetic and genomic applications in common bean. *G3 Genes, Genomes, Genet*. 2015;5(11):2285–90.
51. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
52. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol*. 2005;14(8):2611–20.
53. Earl D, Von Holdt B. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Genet Resour*. 2012;(4):359–61.
54. Mamidi S, Chikara S, Goos RJ, Hyten DL, Annam D, Moghaddam SM, et al. Genome-Wide Association Analysis Identifies Candidate Genes Associated with Iron Deficiency Chlorosis in Soybean. *Plant Genome*. 2011;4(3):154–64.
55. Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, et al. Ensembl Genomes 2020—enabling non-vertebrate genomic research. *Nucleic Acids Res*. 2020;48(D1):D689–95. <https://academic.oup.com/nar/article/48/D1/D689/5584694>

# Figures

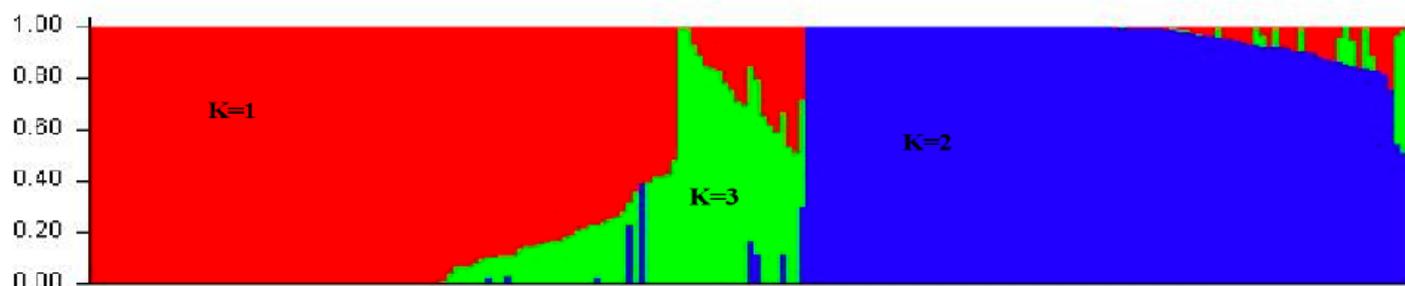


Figure 1

Population structure of 206 common bean genotypes with 5052 SNP markers, whereby red represent Mesoamerican gene pool (K=1), blue represent Andean gene pool (K=2) and green represent admixture of the two pools (K=3).

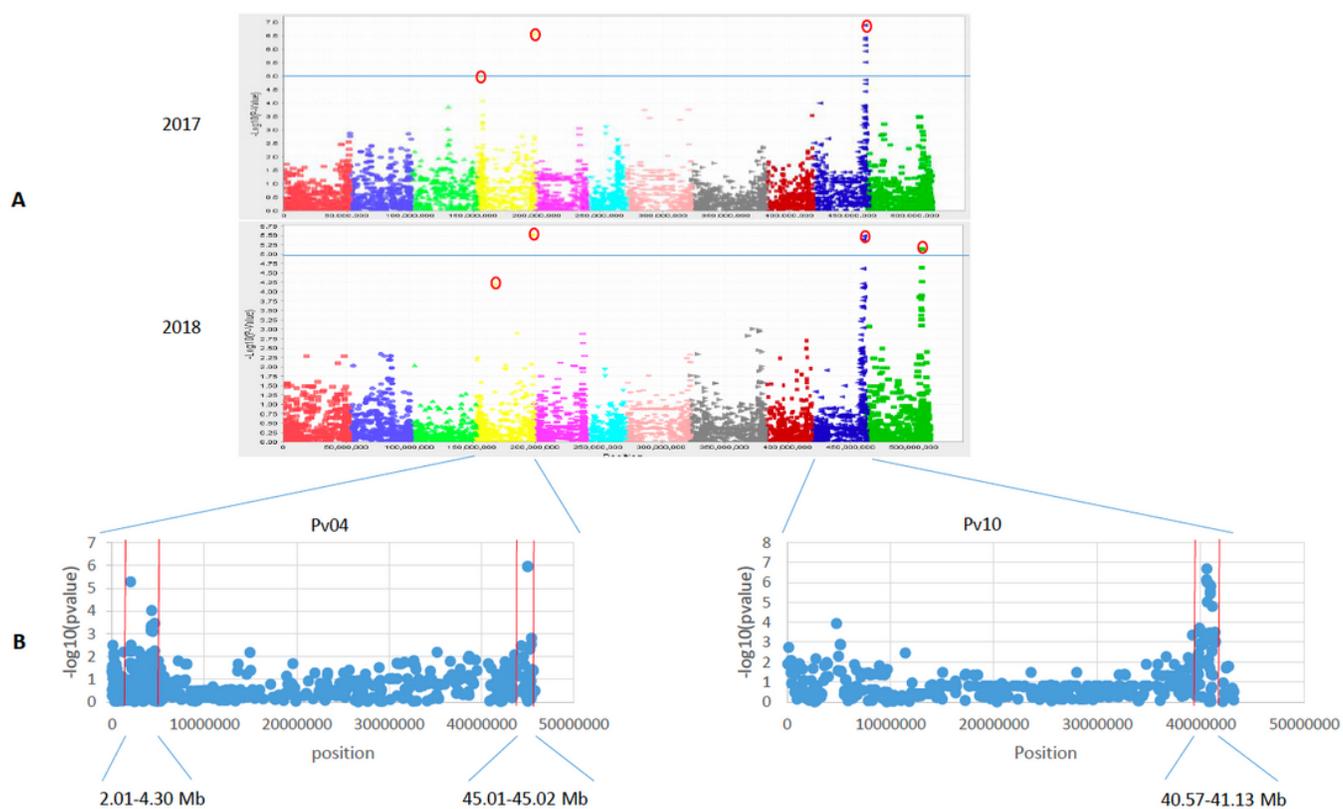


Figure 2

A. Manhattan plots of two seasons for PM disease in 206 collection of bean genotypes using 5052 SNP markers. The horizontal line is the cut off for significance and small red round indicated the most significant markers . B. Distribution of SNPs on Pv 0 4 and Pv10 and the range of SNP peaks.

## Supplementary Files

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

- [Additionalfile1.pdf](#)