

Fine-mapping and evolutionary history of R-BPMV, a dominant resistance gene to Bean pod mottle virus in *Phaseolus vulgaris* L.

Chouaïb MEZIADI

corebiogenesis

Juan-Camilo ALVAREZ-DIAZ

INRAE UMR1403: Institut des Sciences des Plantes de Paris Saclay

Vincent THAREAU

onlogeapied

Ariane GRATIAS

Universite Paris-Saclay

William MARANDE

INRAe

Alvaro SOLER-GARZON

USDA

Phillip N. MIKLAS

USDA

Stéphanie PFLIEGER

Université Paris Cité: Université Paris Cité

Valerie Geffroy (✉ valerie.geffroy@u-psud.fr)

INRAE UMR1403: Institut des Sciences des Plantes de Paris Saclay <https://orcid.org/0000-0003-2210-3636>

Research Article

Keywords: Bean pod mottle virus, *Phaseolus vulgaris*, soybean, *Medicago truncatula*, BAC library, fine mapping, NLR, disease resistance gene clusters, I cluster, legumes

Posted Date: October 9th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3358911/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Theoretical and Applied Genetics on December 13th, 2023. See the published version at <https://doi.org/10.1007/s00122-023-04513-9>.

Fine-mapping and evolutionary history of *R-BPMV*, a dominant resistance gene to *Bean pod mottle virus* in *Phaseolus vulgaris* L.

Chouaïb MEZIADI ^{1,2*}, Juan-Camilo ALVAREZ-DIAZ ^{1,2*}, Vincent THAREAU^{1,2}, Ariane GRATIAS^{1,2}, William MARANDE³, Alvaro SOLER-GARZON⁴, Phillip N. MIKLAS⁵, Stéphanie PFLIEGER^{1,2} and Valérie GEFFROY^{1,2**}

1. Université Paris-Saclay, CNRS, INRAE, Université Evry, Institute of Plant Sciences Paris-Saclay (IPS2), 91190, Gif sur Yvette, France.

2. Université Paris-Cité, CNRS, INRAE, Institute of Plant Sciences Paris-Saclay (IPS2), 91190, Gif sur Yvette, France

3. Washington State Univ, Irrigated Agr Res & Extens Ctr, Prosser, WA USA

4. USDA ARS, Grain Legume Genet & Physiol Res Unit, Prosser, WA USA

5. CNRVG, INRA, F-31326 Castanet Tolosan, France

* These authors contributed equally to this work.

** Correspondence: valerie.geffroy@ips2.universite-paris-saclay.fr. Institute of Plant Sciences Paris-Saclay, Rue Noetzlin, 91405 Orsay, France. Telephone: 33-1-69-15-33-65

Key message: *R-BPMV* is located within a recently expanded TNL cluster in the *Phaseolus* genus with suppressed recombination and known for resistance to multiple pathogens including potyviruses controlled by the *I* gene.

Abstract

Bean pod mottle virus (BPMV) is a comovirus that infects common bean and legumes in general. BPMV is distributed throughout the world and is a major threat on soybean, a closely related species of common bean. In common bean, BAT93 was reported to carry the *R-BPMV* resistance gene conferring resistance to BPMV and linked with the *I* resistance gene. To fine map *R-BPMV*, 182 recombinant inbred lines (RILs) derived from the cross BAT93 x JaloEEP558 were genotyped with polymerase-chain (PCR)-based markers developed using genome assemblies from G19833 and BAT93, as well as BAT93 BAC clone sequences. Analysis of RILs carrying key recombination events positioned *R-BPMV* to a target region containing at least 16 TIR-NB-LRR (TNL) sequences in BAT93. Because the *I* cluster presents a suppression of recombination and a large number of repeated sequences, none of the 16 TNLs could be excluded as *R-BPMV* candidate gene. The evolutionary history of the TNLs for the *I* cluster were reconstructed using microsynteny and phylogenetic analyses within the legume family. A single *I* TNL was present in *Medicago truncatula* and lost in soybean, mirroring the absence of complete BPMV resistance in soybean. Amplification of TNLs in the *I* cluster predates the divergence of the *Phaseolus* species, in agreement with the emergence of *R-BPMV* before the separation of the common bean wild centers of diversity. This analysis provides PCR-based markers useful in marker-assisted selection (MAS) and laid the foundation for cloning of *R-BPMV* resistance gene in order to transfer the resistance into soybean.

Key words

Bean pod mottle virus, *Phaseolus vulgaris*, soybean, *Medicago truncatula*, BAC library, fine mapping, NLR, disease resistance gene clusters, *I* cluster, legumes

Introduction

Common bean (*Phaseolus vulgaris* L.) is the most important grain legume for direct human consumption worldwide, and the most consumed grain legume in Europe with the demand increasing over the last years (Bellucci et al. 2021). In developing countries (Latin America and East Africa), dry beans represent a crucial source of protein, carbohydrates, vitamins and minerals for low income farmers (Broughton et al. 2003; Foyer et al. 2016) (<http://faostat.fao.org/>). Common bean is a diploid ($2n=2x=22$), annual and predominantly self-pollinating species with a relatively small genome (~600 Mb) (Arumuganathan and Earle 1991). This legume (*Fabaceae*) species benefits from well described centers of origin and diversification which makes this plant an attractive model species for evolutionary studies (Bitocchi et al. 2017). The wild forms of common bean grow from northern Mexico to north-western Argentina (Toro et al. 1990), and are organized in three eco-geographic centers of diversity: two major (the Andean and Mesoamerican) and a third minor one (wild populations that grow in northern Peru and Ecuador). About 8000 years ago, two independent domestication events spawned two gene pools of cultivated bean, namely the Andean and Mesoamerican, while no domestication occurred in the northern Peru/Ecuador center of diversity (Shii et al. 1980; Gepts and Bliss 1985; Koinange and Gepts 1992). In addition, the *Phaseolus* genus is well characterized with five (*Phaseolus vulgaris*, *P. coccineus*, *P. dumosus*, *P. acutifolius*, and *P. lunatus*) of its ~70 species that have been domesticated (Bitocchi et al. 2017). The divergence among these five *Phaseolus* species is relatively recent (~2–4 Ma ago; Delgado-Salinas et al., 2006). In common bean, several genome assemblies are now available. In particular, the Andean landrace genotype G19833 was sequenced by Schmutz et al. (2014) providing a great opportunity for the development of molecular markers linked to disease resistance genes in common bean (Meziadi et al. 2016). In addition, the release of the genome assembly of the Mesoamerican cultivated line BAT93 gave a boost, not only for molecular breeding programs but also for understanding the evolution process among the Andean gene pool and Mesoamerican gene pool (Vlasova et al. 2016).

In the past three decades, marker assisted selection (MAS) has been largely used in common bean to improve various traits including pathogen resistance (Miklas et al. 2006; Ferreira et al. 2012; Kelly and Bornowski 2018). The concept consists in introducing specific resistance (*R*) genes into one genotype and/or to pyramid *R* genes to achieve durable resistance

(Miklas et al. 2006; Ferreira et al. 2012; Mundt 2014). In that context, identification of robust genetic resistance to infectious pathogens is a major challenge.

To fight against pathogen attacks, plants have evolved multiple immune receptors for the activation of defense responses (Dodds and Rathjen 2010; Dangl et al. 2013). Nucleotide-binding domain leucine-rich repeat (NLR) proteins are intracellular immune receptors that monitor pathogen invasion *via* direct or indirect sensing of effectors derived from pathogens (Dangl et al. 2013). After recognition, activated NLRs initiate an immune signaling leading to resistance, called effector-triggered immunity (ETI) (Dangl et al. 2013). These NLR genes correspond to the major dominant resistant (*R*) genes described by Harold Henry Flor in the 1940s in the context of the gene-for-gene resistance (Flor 1956). NLR R proteins can recognize the presence of various disease-causing pathogens including bacteria, oomycetes, fungi, nematodes, insects, and viruses (Cesari et al. 2014; Kourelis and van der Hoorn 2018; Ngou et al. 2022). Among this prevalent class of *R* genes, two subclasses have been identified based on their N-terminal domain: the Coiled-Coil NLRs (CNLs) and the Toll Interleukin-1 receptor NLR (TNL) (Meyers et al. 1999, 2003; Tamborski and Krasileva 2020). Genomic analysis have demonstrated that NLR sequences are abundant and often physically clustered in many plant genomes (Michelmore et al. 2013; Dangl et al. 2013). These clusters formed by gene duplication, unequal crossing-over, and ectopic recombination provide a reservoir of genetic diversity for generating new resistance specificities (David et al. 2009; Tamborski and Krasileva 2020; Barragan and Weigel 2021; Märkle et al. 2022). Like other crop species, NLR genes in common bean are often organized in clusters of tightly linked genes such as the *B4* CNL cluster located at proximal end of chromosome 4 (Pv04) (Ferrier-Cana et al. 2003, 2005; Geffroy et al. 2009; David et al. 2009), the *Co-2* CNL cluster at distal end of Pv11 (Creusot et al. 1999; Innes et al. 2008; Chen et al. 2010) and the *I* cluster at distal end of Pv02 containing a cluster of TNL sequences (Vallejos et al. 2006).

In common bean, the *I* resistance cluster is a multi-parasitic resistance cluster (Meziadi et al. 2016) that controls resistance against 11 viruses [BCMV (*Bean common mosaic virus*), BCMNV (*Bean common mosaic necrosis virus*), ZYMV (*Zucchini yellow mosaic virus*), PWV-K (*Passionfruit woodiness virus-K*), WMV (*Watermelon mosaic virus-2*), CabMV (*Cowpea aphid-borne mosaic virus*), ThPV (*Thailand passiflora mosaic virus*), SMV (*Soybean mosaic virus*), BYMV (*Bean yellow mosaic virus*), CLYVV (*Clover yellow vein virus*) and BSMV (*Bean severe mosaic virus*)] as well as against the bacteria *Pseudomonas syringae* pv. *phaseolicola* and the fungus *Colletotrichum lindemuthianum* (Fisher and Kyle 1994, 1996; Crute and Pink 1996; Morales and Singh 1997; Geffroy et al. 2008; Bello et al. 2014; Tock et

al. 2017). Recently, the dominant *R-BPMV* gene, conferring resistance against *Bean pod mottle virus* (BPMV, genus *Comovirus*, family *Comoviridae*), was also linked to the *I* cluster (Pflieger et al. 2014; Meziadi et al. 2021). Like other comoviruses, BPMV has a bipartite-positive strand RNA genome consisting of RNA1 and RNA2, which are separately encapsidated in isomeric particles (Ghabrial 1982). BPMV is distributed throughout the world and is a major threat on soybean (*Glycine max*), especially in the United States where all commercially available cultivars of soybean are susceptible to BPMV (Giesler et al. 2002). At the molecular level, the *I* cluster co-localized with a cluster of 7 TNL in G19833 (Schmutz et al. 2014).

Evolutionary relationships between two genotypes of the same and/or different related species is a key tool for comparative genomics. Related species have a common ancestor but have undergone independent evolution process. As a result of sequencing efforts, genome assemblies are available in various legume species including *Medicago truncatula*, *G. max*, and *P. vulgaris*, as well as *P. lunatus* and *P. acutifolius* (Schmutz et al. 2010, 2014; Young et al. 2011; Garcia et al. 2021; Moghaddam et al. 2021). Consequently, the legume family is ideally suited for comparative phylogenomic approaches, in which phylogenetic inference is combined with structural genomic analyses (Ammiraju et al. 2008). From an applied point of view, comparative genomics can be a powerful approach to gene discovery in related species (McClellan et al. 2010). Because common bean is phylogenetically close to soybean, understanding evolution of *R* genes in these two related species can offer a better understanding of genome dynamics and NLR evolution.

Here, we present the fine mapping of *R-BPMV*, a dominant *R* gene conferring resistance against BPMV using a population of recombinant inbred lines (RILs) derived from a cross between BAT93 (R) × JaloEEP558 (S). PCR-based markers developed in this study using the two common bean genome assemblies of G19833 and BAT93, constitute the raw material to develop markers for MAS for the *R-BPMV* gene as well as for the other linked resistance genes at the *I* cluster. This study was conducted in a favorable context for understanding the evolution of resistance genes on a short time scale (between two *P. vulgaris* genotypes of the two different gene pools) and at longer time scale (between genus *Phaseolus*, *Glycine* and *Medicago*). Thanks to the comparison of the *I* cluster corresponding regions in *P. vulgaris*, *G. max* and *M. truncatula*, we have shown that the *I* TNL cluster has amplified recently in the *Phaseolus* genus as a result of tandem duplications and presents copy number variations (CNV) between common bean genotypes.

Materials and methods

Plant material and growth conditions

An F11 population of 182 RILs, derived from the cross between the Andean landrace JaloEEP558 and the Mesoamerican breeding line BAT93, was used to map the *R-BPMV* resistance gene. This population is composed of 77 RILs used to set up the integrated linkage map of common bean (Freyre et al. 1998) and 105 additional RILs developed in Orsay (Chen et al. 2010). G19833, a CIAT germplasm accession, is a landrace ‘Chaucha Chuga’ originally collected in Peru that belongs to the Andean gene pool of common bean. In addition to BAT93 and G19833, a set of 17 cultivated common bean genotypes (Table 1) from our Orsay collection as well as 8 wild beans (Table 2) from CIAT (Centro Internacional de Agricultura Tropical; Colombia) were used for inoculation with BPMV isolate IADi-1. Seedlings were cultivated in plastic pots filled with moist vermiculite at 23 °C, 75% relative humidity, with 8 h dark and 16 h light during 10–12 days (the stage at which primary leaves are fully expanded).

Viral pathogenicity test

The BPMV strains (BPMV-GFP; BPMV isolate IADi-1) used in this study were previously described in Pflieger et al. (2014). The BCMNV NL-3 strain from Bello et al. (2014) which traces back to Drijfhout’s (1978) original collection was used. Inoculations with BCMNV NL-3 strain were performed in common bean land race JaloEEP558. Mosaic symptoms were observed in primary inoculated leaves at 7-10 days after inoculation. These leaves were harvested and used as inoculum. For the inoculation method, infected plants with BPMV-GFP or BCMNV NL-3 strain (fresh leaf or frozen leaf conserved at -80°C) were ground in a mortar with 50 mM potassium phosphate buffer (pH 7) to make leaf sap. Mechanical inoculations were performed by rubbing of the primary leaves of plantlets (10- to 12-days old) using carborundum as an abrasive. After viral inoculation, plants were placed in a growth chamber at 20°C under a 16-h light/8-h dark cycle under a humidity of 70% and watered with nutrient solution (Pflieger et al. 2015). Seven days after inoculation with BPMV-GFP, plants without GFP fluorescence and visible symptoms or showing necrotic pinpoint on inoculated leaf were scored as resistant, while plants with fluorescent local lesions on the inoculated leaf were scored as susceptible. This is slightly different from Pflieger et al. (2014) where plants were scored as susceptible only when GFP fluorescence was observed in systemic leaves. For NL-3 BCMNV strain, 7-10 days after inoculation, plants showing top necrosis were scored resistant and those showing systemic

mosaic were scored susceptible. Two replicates of all experiments were carried out. The resistance/susceptible phenotype was assessed on at least 12 seedlings for each RIL or bean genotype.

GFP imaging

GFP expression in aerial organs was examined using a UV lamp (High intensity 100-Watt long-wave UV lamp; UVP, USA) and photographed using a Canon PowerShot A2300 digital camera.

DNA extraction

Nuclear DNA was isolated as described in Geffroy et al. (1998). In brief, 2 g of leaves were ground in liquid nitrogen. The powdered tissue was transferred into 50 mL of cold buffer 1 (10 mM Tris pH 7.5; 15 mM KCl; 15 mM NaCl; 0.15 mM spermine; 0.5 mM spermidine; 2 mM EDTA; 0.5% triton-X100 and 4.5% glucose) and filtered through blutex. The filtrate was centrifuged for 10 min at 2,600 rpm and 4°C and the supernatant was discarded. The pellet was rinsed with 5 mL of buffer-1 three times and was finally resuspended into 1 mL of distilled water. Afterwards, a CTAB extraction procedure was performed as described in Doyle and Doyle (1991) except that the chloroform–isoamyl alcohol extraction step was replaced by a chloroform–octanol (24:1) extraction step.

PCR-based marker design and RILs genotyping

PCR-based markers were designed using G19833 v1.0 (<https://data.legumeinfo.org/Phaseolus/vulgaris/genomes/G19833.gnm1.zBnF/>) and BAT93 v10 (<https://denovo.cnag.cat/bean>) genome assemblies using Primer3 version 4.0 software (<http://frodo.wi.mit.edu/primer3/>; Rozen and Skaletsky 2000) with default parameters. Twenty-eight primer pairs were designed (listed in Supplementary Tables S2 and S3) mainly in intergenic regions because they usually present higher frequencies of polymorphism than coding regions. Primer pairs were designed to amplify about 0.6 to 1 kb with a GC percent of ~50% and were selected to be locus specific. Primer specificity in the *P. vulgaris* genome was checked using custom Perl scripts based on the SPADS program (Thareau et al. 2003). All the primers were synthesized by Eurogentec Company (Liege, Belgium). These primer pairs were experimentally tested for amplification and polymorphism between the parental lines BAT93 and JaloEEP558. PCR reactions were performed using Go Taq DNA Polymerase kit (Promega, Charbonnières-les-Bains, France) using a standard PCR program (start: 94°C for 5 min; amplification: 35 cycles of 94°C for 30 s, 59–65°C for 45 s, 72°C for 1–2 min; termination for: 72°C for 5 min). PCR products were resolved on 1% agarose gels containing ethidium bromide, run in 0.5X TBE buffer

and visualized under UV light. Primer pairs revealing polymorphism between BAT93 and JaloEEP558 were used to genotype the 182 RILs following the same PCR protocol as described above.

Genetic mapping of the *R-BPMV* gene

A Chi-squared (χ^2) test was used to evaluate the goodness of fit of observed and expected segregation ratios. The MAPMAKER software version 3.0 (Lander et al. 1987) was used to map the *R-BPMV* resistance gene on the integrated linkage map of common bean using the set of 142 markers as described in Geffroy et al. (2008). Linkage groups were established with a LOD threshold of 3.0 and a maximum recombination fraction of 0.3. The position of *R-BPMV* on LG B2 was estimated with a LOD threshold of 2.0 based on the “try” function. Map distances were estimated by the Kosambi mapping function (Kosambi 1943).

BAC library construction, screening, PacBio RS II sequencing, and assembly

Bacterial artificial chromosome (BAC) cloning and high-throughput sequencing were used to isolate genomic regions of *P. vulgaris*, genotype BAT93. High molecular weight (HMW) DNA was prepared from young leaves using the method described by Peterson *et al.* (2000) and modified as described in Gonthier et al. (2010). In brief, twenty grams of frozen leaf tissue was ground to powder in liquid nitrogen with a mortar and pestle and then used to prepare megabase-size DNA embedded in agarose plugs. Embedded HMW DNA was partially digested with *Hind*III (New England Biolabs, Ipswich, MA, USA), subjected to two size selection steps by pulsed-field electrophoresis, using a BioRad CHEF Mapper system (Bio-Rad), and ligated to pIndigoBAC-5 *Hind*III-Cloning Ready vector (Epicentre Biotechnologies, Madison, WI, USA). Pulsed-field migration programs, electrophoresis buffer and ligation desalting conditions were performed according to Chalhoub et al. (2004). The BAT93 BAC library is composed of 18 432 clones with a mean insert size of 143 kb and represents 4,2 genome equivalents. All the clones were spotted on high-density colony filters and screened by hybridization with 5 radiolabelled probes (Supplementary Table S1).

High-density colony filters were prepared with the genomic BAT93 BAC library using a robotic workstation QPix2 XT (Molecular Devices, Berkshire, UK). BAC clones were spotted in duplicate using a 4 × 4 pattern onto 22 × 22 cm Immobilon-Ny+ filters (Molecular Devices). The whole BAC library was spotted in duplicate on one filter. After incubation at 37°C for 17 h, filters were processed as follows: (1) denaturation on Whatman paper soaked with a solution of 0.5 M NaOH and 1.5 M NaCl for 4 min at room temperature, and for 10 min at 100°C, (2)

neutralization on Whatman paper soaked with 1 M TrisHCl pH 7.4, and 1.5 M NaCl for 10 min, incubation in a solution of 0.25 mg mL⁻¹ proteinase K (Sigma Aldrich) for 45 min at 37°C, baking for 45 min at 80°C, and (3) fixation by UV on a Biolink 254 nm crosslinker (Thermo Fischer Scientific, Waltham, MA, USA) with an energy of 120 000 µJoules. Radiolabelling of probes and hybridization of the filters were performed as described in Gonthier et al. (2010). Hybridized filters were imaged with a Storm 860 PhosphorImager (GE Healthcare), and analyses were performed using the HDFR software (Incogen, Williamsburg, VA, USA). Positive BAC clones detected by hybridization were validated individually by PCR amplification using the primer pairs used for probes synthesis and by visualization of PCR products after agarose gel electrophoresis. BAC ends were sequenced using Sanger sequencing and primers located in the plasmids.

Four BAC clones (29P20; 29I04; 11D15; 46M20), were chosen for sequencing. A pooled library combining 2 µg of each BAC clone was generated using the standard Pacific Biosciences library preparation protocol for 8–12 kb libraries. This library was sequenced in one PacBio RS II SMRT Cell using the P6 polymerase in combination with the C4 chemistry (sequencing service following the standard operating procedures was provided by IGM Genomics Center).

Assembly of the PacBio RS II reads was performed following the Hierarchical Genome Assembly Process (HGAP) workflow. The SMRT[®] Analysis (v.2.2.0) software suite was used for HGAP implementation (<https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP>). The reads were first aligned using Blast (Blast on Pacific Biosciences repository website against the *Escherichia coli* str. K12 substr. DH10B, complete genome). Identified *E. coli* reads and low quality reads (read quality < 0.80 and read length < 500 bp) were removed from the data used for the BAC clone sequence assembly. Vector sequences were removed before the final assembly process. The correspondence between the BAC assembly and the BAC clones was done by matching their BAC-end sequences on the ends of assembled sequences with BLAST. Genbank accession numbers of BAC contigs UTG0 and UTG1 are XXXX and XXXX, respectively.

BAC sequence annotation

BAC sequences were annotated using an automatic annotation pipeline described in David et al. (2009). The pipeline uses a combination of gene-finding and sequence homology programs with known genes and proteins: the two *ab initio* gene prediction programs FGENESH (Burset and Guigó 1996) and GeneMarkhmm (Lukashin 1998), and BLAST (Altschul 1997) for the analyses

against the GenBank nonredundant database. All this information was imported into the annotation platform Artemis (Rutherford et al. 2000) for further manual analysis and compared with the annotation provided in Phytozome. Annotation of NLR genes was performed by using the program NLR annotator (v1) (Steuernagel et al. 2020) and MEME (v4.9.1) (Bailey et al. 2015).

Synteny analysis

Genome sequences and respective gene annotations were downloaded for *P. vulgaris* G19833 v1.0 (<https://data.legumeinfo.org/Phaseolus/vulgaris/genomes/G19833.gnm1.zBnF/>), *P. vulgaris* BAT93 v10 (<https://denovo.cnag.cat/bean>), soybean (*Glycine max*) (https://phytozome-next.jgi.doe.gov/info/Gmax_Wm82_a2_v1), *Medicago truncatula* (https://phytozome-next.jgi.doe.gov/info/Mtruncatula_Mt4_0v1), *Phaseolus acutifolius* Frijol Bayo (https://phytozome-next.jgi.doe.gov/info/Pacutifolius_v1_0), *Phaseolus acutifolius* W6 15578 (https://phytozome-next.jgi.doe.gov/info/PacutifoliusWLD_v2_0) and *Phaseolus lunatus* (https://phytozome-next.jgi.doe.gov/info/Plunatus_V1). Annotation of NLR genes was performed for all genome assemblies by using the program NLR annotator (v1) (Steuernagel et al. 2020) and MEME (v4.9.1) (Bailey et al. 2015). SyMAP v4.0 (Soderlund et al. 2011) was used to identify syntenic blocks between the genomes, using default parameters.

Phylogenetic analysis of TNL sequences

For TNL phylogenetic analysis, we used TNL sequences from the reference genome assembly G19833 v1.0 (Schmutz et al. 2014), TNL sequences from BAT93 BAC contigs corresponding to the *I* cluster (UTG0 and UTG1; this study), and a representative subset of TNL sequences from *M. truncatula*. A multiple alignment of amino acid sequences from the NB portion of TNLs (from the P-loop to the MHD motif) was performed with MAFFT version 6 using the L-INS-I strategy (Kato 2005) with default parameters. Alignments were then subjected to maximum likelihood (ML) analysis using the JTT model as implemented in PhyML (Guindon and Gascuel 2003). Relative support for clades was assessed with 100 bootstrap replicates.

Results

Identification of PCR-based markers flanking the *R-BPMV* resistance gene

The *R-BPMV* gene is present in genotype BAT93 of *P. vulgaris* and confers resistance against BPMV (Pflieger et al. 2014). This gene was previously mapped at distal end of LG Pv02, between markers DROS3b and D1287, at ~6.9 cM and ~14 cM, respectively, using a set of 77 RILs derived from a cross between BAT93 × JaloEEP558 (Pflieger et al. 2014) (Fig. 1A). To identify closer flanking markers, the PCR-based markers IG19M100a and IG19M500b were designed using the reference genome assembly from G19833 v1.0 and found to be polymorphic between BAT93 and JaloEEP558 (presence/absence polymorphism). Using the same set of 77 RILs population, IG19M100a and IG19M500b were mapped at 7.5 cM and 5.8 cM from *R-BPMV*, respectively (Fig. 1A). In G19833 v1.0 genome assembly, the physical distance between markers IG19M100a and IG19M500b corresponds to an ~2 Mb interval, containing a cluster of TNL-encoding genes (Fig. 1).

Fine mapping of the *R-BPMV* gene

To increase resolution for ‘fine’ mapping the *R-BPMV* gene, the size of the RILs mapping population was increased from 77 to 182 by adding 105 additional RILs from the same BAT93 × JaloEEP558 cross (Chen et al. 2010). These 105 additional RILs were inoculated with the BPMV-GFP construct derived from BPMV isolate IA-Di1 (Zhang et al. 2010). Resistant and susceptible phenotypes were scored visually at 7 to 10 days post-inoculation (dpi). The ratio of susceptible (95 RILs) to resistant (87 RILs) of the total 182 RILs was evaluated using Chi-squared tests, and the observed segregation ratio fitted the expected 1:1 ratio of susceptible to resistant plants ($\chi^2 = 0.35$, $p = 0.6$) confirming that a single gene is segregating.

The design of additional PCR-based markers, based on G19833 v1.0 genome assembly, in combination with the 182 RILs population narrowed the target region. *R-BPMV* was mapped between markers IG19M201b and IG19M401b at 1.8 and 1.6 cM, respectively (Fig. 1A). The physical distance between IG19M201b and IG19M401b, according to G19833 genome assembly, is 169 kb (Fig. 1B).

To identify the *R-BPMV* corresponding region in BAT93 v10 genome assembly (Vlasova et al. 2016), IG19M201b and IG19M401b markers were used as queries in BLASTN analysis against BAT93 v10 genome assembly. Two scaffolds were identified, scaffold 00018 (1.57 Mb) and scaffold 00140 (791 kb) (Fig. 2). A high level of synteny was identified between scaffolds 00018 and 00140, nearly depleted of TNL, and the corresponding region in G19833

genome assembly (Fig. 2). In BAT93, the size of the gaps between these two scaffolds 00018 and 00140 is unknown (Fig. 2).

Because the BAT93 genome assembly is fragmented in the *R-BPMV* region, a BAT93 BAC library was constructed and screened with PCR-based markers derived from BAT93 scaffolds 00018, 00140 (Supplemental data Table S1). This led to identify 4 positive BAT93 BAC clones presenting a size ranging from 100 to 180 kb. After PCR amplification and sequencing, these BAC clones were organized into two non-overlapping contigs UTG1 (160 kb) and UTG0 (237 kb) (Fig. 2). Compared to scaffold 00018 and 00140, UTG1 and UTG0 add 77 and 167 kb, respectively (Fig. 2). Manual annotation of UTG1 and UTG0 revealed the presence of sixteen TNL genes all in the same orientation, 5 on UTG1 and 11 on UTG0 (Fig. 2). Thus, the *R-BPMV* target region seems to be much larger in BAT93 harboring 16 TNL compared to G19833 harboring only 7 TNL (Fig. 1C), and this is probably an underestimation in BAT93 since a gap is still present between UTG0 and UGT1.

To restrict the number of *R-BPMV* TNL candidate gene(s), additional polymorphic PCR-based markers were generated using either BAT93 or G19833 sequences and first tested on genomic DNAs of the two parents BAT93 and JaloEEP558. Nineteen markers positioned between IG19M201b and IG19M401b were used to genotype the 3 RILs presenting a recombination event between markers IG19M201b and IG19M401b. Strikingly, co-segregation was observed between the *R-BPMV* resistance phenotype and the 16 markers covering the TNL cluster (Supplementary data Tables S2 and S3; Supplementary data Fig. S1). Therefore, despite many additional markers, *R-BPMV* target region was only slightly reduced between markers M11RB10 and IG19M401b and none of the 16 candidate TNL genes could be excluded as *R-BPMV* candidate.

***R-BPMV* cosegregates with the *I* resistance gene**

R-BPMV was previously linked to the *I* resistance cluster located at distal end of Pv02 (Pflieger et al. 2014) where several TNL genes were annotated in G19833 (Fig. 1C). This cluster confers resistance to viral, bacterial and fungal diseases (Meziadi et al. 2016). One *R* gene described in this locus confers a hypersensitive systemic necrosis resistance response (HR) to the NL-3 strain of BCMNV and is known as the *I* gene (Freyre et al. 1998; Vallejos et al. 2006). To examine whether the resistance against BPMV and HR resistance to BCMNV segregates as a single locus, the population of 182 RILs derived from the cross BAT93 × JaloEEP558 was inoculated with the NL-3 BCMNV strain and phenotyped visually at 7 to 10 dpi. BAT93 which

carries the dominant allele *I*, exhibited HR (systemic necrosis phenotype) within 7 dpi, whereas systemic mosaic (susceptible phenotype) was observed on JaloEEP558 within 10 dpi. The 182 RILs were phenotyped by visual inspection of symptoms and categorized as ‘HR - resistant’ for systemic top necrosis and ‘susceptible’ for systemic mosaic expression, respectively. Perfect co-segregation was observed with the 182 RILs for the *R-BPMV* and *I* gene HR phenotype with no recombinant RILs (Fig. 1A). This result confirms that *R-BPMV* is linked to the ‘*I* gene’ within the same TNL gene cluster and suggests that resistance to BPMV and BCMNV is conferred by either a single gene or by two closely linked genes that segregate as a single locus.

Origin and evolution of the *I* TNL cluster

P. vulgaris and *M. truncatula* are two legume species that diverged ~54 million years ago (Mya) while *P. vulgaris* and *G. max* diverged more recently ~19 Mya (Lavin et al. 2005; Innes et al. 2008; Stefanović et al. 2009). After the *P. vulgaris*/*G. max* divergence, the *G. max* genome has undergone a whole genome duplication event, estimated to have occurred ~10 to 14 Mya (Shoemaker et al. 2006). BPMV has been reported to be a major threat to soybean production in the USA, however, to our knowledge, all commercially available cultivars of soybean are susceptible to BPMV (Giesler et al. 2002). In order to identify in the soybean genome, the region corresponding to the *I* resistance cluster, two “single-copy genes” flanking the common bean *I* cluster (Phvul.002G320700 and Phvul.002G324500) were used for microsynteny analysis (Fig. 3). As expected, two corresponding homeologous regions were identified in the *G. max* genome (McClellan et al. 2010) on chromosomes 5 and 8. These regions present a high level of microsynteny over the ~200 kb studied in soybean, with nearly all the ~thirty genes conserved in order and orientation between *P. vulgaris* and *G. max* (Fig. 3). However, strikingly, no TNL gene was annotated in the two corresponding homeologous regions of *G. max* (Fig. 3).

To investigate the evolution of the TNL sequences present at the *I* cluster on a longer time scale, further microsynteny studies were conducted between *P. vulgaris* and *M. truncatula* (Supplementary data Fig. S2). Two hypotheses were tested: i) TNLs were inserted at the *I* locus after *P. vulgaris* and *G. max* diverged (gain hypothesis) or ii) alternatively TNLs were lost from *G. max* after divergence (loss hypothesis). The same two flanking genes (Phvul.002G320700 and Phvul.002G324500) were used to identify the corresponding region in *M. truncatula* genome. Compared to common bean (291 kb), the region corresponding to the *I* locus is larger in *M. truncatula* (427.5 kb). As observed with *G. max* genome, excluding the TNL, the two syntenic regions are highly colinear between *P. vulgaris* and *M. truncatula*, with 25 out of 30

genes conserved in order and orientation. Interestingly, contrasting with *G. max* genome, one TNL gene (Mt8g105820) was identified in the syntenic region in *M. truncatula* (Supplementary data Fig. S2). Phylogenetic analyses showed that this *M. truncatula* TNL (Mt8g105820) and TNLs belonging to the *I* resistance cluster in common bean (G19833, BAT93) are grouped into one well-supported clade (Fig. 4). This indicates that this unique TNL of *M. truncatula* and the TNLs of *P. vulgaris* from the *I* cluster are orthologous. Furthermore, the absence of TNLs in the region syntenic of the *I* cluster in *G. max* and the presence of one TNL in *M. truncatula* supports the hypothesis of loss of TNLs in *G. max*.

To better date the amplification of the TNL from the *I* cluster, we took advantage of three recently sequenced *Phaseolus* genomes, namely *P. acutifolius* (Frijol Bayo, cultivated), *P. acutifolius* (W6 15578, wild) and *P. lunatus* (G27455) (Moghaddam et al. 2021; Wisser et al. 2021). Like in *G. max* and *M. truncatula*, the syntenic region corresponding to the *I* cluster was clearly identified in both species with many genes conserved in order and orientation (Supplementary data Fig. S3). Interestingly, many TNLs were identified in both *P. acutifolius* (10 TNLs) and *P. lunatus* (6 TNLs), suggesting that the TNL amplification at the *I* cluster predates the divergence of the *Phaseolus* species.

Origin of *R-BPMV* resistance gene

To gain insight into the origin of the *R-BPMV* gene, we used a panel of cultivated and wild common bean genotypes from various geographical origins. Concerning cultivated genotypes, sixteen genotypes from both Mesoamerican and Andean gene pools were selected in addition to BAT93, JaloEPP558 and G19833. These genotypes were infected with BPMV-GFP (Table 1). Ten genotypes belonging to both Mesoamerican and Andean cultivated gene pools were found to be resistant to BPMV (Table 1). Genotyping with MB29530 PCR-based marker, co-segregating with *R-BPMV* (Supplementary data Fig. S1; Supplementary data Table S3), confirms that these 10 genotypes carry the *R-BPMV* gene (Table 1). Concerning the wild accessions, five accessions were found to be resistant to BPMV according to their phenotypes and the presence of *R-BPMV* gene was confirmed by marker MB29530 profile (Table 2). Interestingly, these five wild accessions come from various countries covering the three centers of diversity of common bean namely Andean, Meso-American and Northern Peru/Ecuador centers of diversity. Altogether, this strongly suggests that the emergence of *R-BPMV* gene predates the separation of the wild centers of diversity.

Discussion

In order to identify the molecular basis of the *R-BPMV* resistance gene linked to the *I* cluster in common bean, a fine mapping strategy was carried out on a population of 182 RILs derived from a cross between the Mesoamerican resistant landrace BAT93 and the Andean susceptible landrace JaloEEP558. This study was conducted in a favorable context since *i*) RIL populations are known to be as informative as F2 populations for small genetic distances (Fisher, 1937), *ii*) the size of the RIL population combined with the high level of polymorphism between the Andean and the Mesoamerican parents should have allowed a good mapping resolution, *iii*) the distal location of the *I* cluster was also a positive point, since distal regions are highly recombinant as compared to centromeric regions (Schmutz et al. 2010, 2014). Such a configuration was present for the successful cloning of the *Co-x* anthracnose gene (Richard et al. 2014, 2021). Concerning the *I* cluster, the present fine mapping strategy was confounded by suppression of recombination combined with the presence of repeated sequences. However, results obtained in this study provide important information on both the structural organization and the evolution history of the *I* cluster paving the way for future research.

The two common bean genome assemblies for BAT93 (Mesoamerican genotype carrying the *R-BPMV* gene) and G19833 (Andean gene pool) constituted important resources to design markers for the fine mapping of *R-BPMV* (Schmutz et al. 2014; Vlasova et al. 2016). Nevertheless, the poor quality of BAT93 genome assembly in the *I* cluster region, fragmented in badly assembled small contigs (data not shown), led us to develop a BAT93 BAC library, a valuable resource not only for the present study but also for future studies for the bean community. Concerning the *I* cluster, despite a gap still present in the BAT93 BAC contigs, the BAC library was crucial since nearly all the 16 BAT93 TNL sequences were identified from these BAC clones, and not from BAT93 genome assembly (Vlasova et al. 2016), demonstrating the complexity of this TNL-containing region. PCR-based markers were designed based on polymorphism identified in the region of interest between G19833 and BAT93. The reference genotype G19833 belongs to the same gene pool as the susceptible parent JaloEEP558, allowing the quick development of successful polymorphic markers. However, locus specific markers were difficult to identify. Altogether, the difficulty to design locus-specific markers and the poor quality of BAT93 genome assembly (Vlasova et al. 2016) in that region testifies to the highly repeated nature of the *I* cluster. The subtelomeric *Khipu* satellite repeats interleaved between CNL at the B4 and Co-2 clusters is not present between the TNL genes of the *I* cluster

(David et al. 2009; Richard et al. 2013; Chen et al. 2018). Consequently, additional work will be needed to characterize these intergenic repeats. Nevertheless, the 23 PCR-based markers developed in the present study constitute the raw material for implementation of marker assisted selection (MAS) in common bean breeding programs (Supplementary data Tables S2 and S3). Indeed, numerous disease resistance genes have been linked to the *I* cluster (Meziadi et al. 2016).

Taking advantage of the genome assembly resources available for *P. vulgaris*, *P. acutifolius*, *P. lunatus*, *G. max*, and *M. truncatula* in a comparative genomic study centered on the *I* cluster, we gained insights into the evolution of this TNL cluster on both short time scale (between two *P. vulgaris* genotypes) and long time scale (between genus *Phaseolus*, *Glycine* and *Medicago*). Microsynteny analysis of the region containing the *I* cluster revealed clear syntenic regions among these legume species. Even though the most recent common ancestor between *P. vulgaris*, *G. max*, and *M. truncatula* is dated 54 Mya (Lavin et al. 2005; Cronk et al. 2006), a high level of microsynteny was observed between the three species for this genomic region (Fig. 3 and Supplemental data Fig. S2), confirming previous synteny analysis performed on these legume species (David et al. 2009; McClean et al. 2010). Strikingly, this collinearity was restricted to non-TNL-encoding genes. Indeed, TNL sequences were present in both *P. vulgaris* (7 TNL in G19833) and *M. truncatula* (1 TNL; Mt8g105820) syntenic regions but they were completely absent in the two corresponding *G. max* syntenic regions. Phylogenetic analysis showed that Mt8g105820 is orthologous to the common bean *I* TNLs suggesting that a single TNL has undergone tandem duplications in *P. vulgaris* but not in *M. truncatula* and was lost in *G. max*. The fact that they are all in the same orientation on the *I* cluster (Fig. 2) suggests the TNL cluster derives from tandem duplications probably resulting from unequal crossing-over events (Barragan & Weigel, 2021; Choi et al., 2016). The identification of the *I* TNL cluster in both *P. lunatus* and *P. acutifolius* allowed us to date more precisely this TNL amplification, predating the divergence of the *Phaseolus* species (~2–4 Ma ago; Delgado-Salinas, Bibler, and Lavin 2006). The identification of the *R-BPMV* resistance specificity in wild accessions from the three centers of diversity allowed us to conclude that *R-BPMV* gene emerged before the separation of the wild centers of diversity, before domestication. In legumes, more than three-quarters of NLR expanded by local tandem duplications producing closely related NLR genes often located at the same cluster (Shao et al. 2014). Ectopic duplications, producing closely phylogenetically related NLRs genes located distantly on a chromosome or on different chromosomes (David et al. 2009; Shao et al. 2014), contribute approximately to less than 20% of new NLR genes loci (Shao et al. 2014). This is probably the case of the TNL

gene 07G160900 located on the chromosome 07 of *P. vulgaris* (Schmutz et al. 2014) and found in the same clade as TNLs of the *I* cluster (Fig. 4).

In common bean, we identified important copy number variation (CNV) of TNLs in the *I* cluster with 7 TNLs in G19833 and at least 16 observed in BAT93, which is an underestimation since a gap is still present in the BAC contig. This is consistent with previous RFLP analysis where ~24 TNLs were identified in Sprite and ~12 in Calima (Vallejos et al. 2006). Structural heterozygosity at NLR clusters has been correlated with coldspot of recombination (Choi et al. 2016). In our study, no recombination event could exclude any of the 16 TNLs as *R-BPMV* candidate gene. Since recombination during meiosis is known to be reduced by structural differences (Barragan and Weigel 2021), the suppression of recombination observed at the *I* TNL cluster in the cross between BAT93 x JaloEEP558 is probably the result of TNL CNV in this genomic region between BAT93 and JaloEEP558. Sequencing of the *I* cluster in JaloEEP558 will be conducted to further address this question. Interestingly, suppression of recombination at the *I* cluster has already been reported during attempt to clone the *I* gene (Vallejos et al. 2006) or the *By-2* gene conferring resistance against BYMV and CLYVV (Hart and Griffiths 2015), despite the use of large F2 populations of 3056 and 552 individuals, respectively.

Since the discovery of resistance against BCMV conferred by the *I* gene in the early 1930s by Ralph Corbett (Pierce, 1934), resistance to eleven other viruses were linked to the *I* gene for a total of resistance to 10 potyviruses and 2 comoviruses. *R-BPMV* is the second *R* gene conferring resistance to a comovirus that was linked to the *I* gene after the *R* gene against CpSMV reported by Morales and Singh, (1997). Despite the difference between comoviruses and potyviruses, the two groups belong to the same taxonomical family (de Haan et al. 1991). Consequently, it is possible that a single TNL confers resistance against these various viruses. In our study *R-BPMV* co-segregates perfectly with the *I* resistance gene, so we cannot distinguish whether a single TNL or two different TNL confer resistance to these two viruses. Indeed, the distance of 0.7 cM previously reported between *I* and *R-BPMV* gene (Pflieger et al. 2014) was due to a missing data instead of a true recombination event.

R-BPMV is a gene of great interest from an agronomic point of view. BPMV has been reported to be a major threat to soybean production in the USA (Giesler et al. 2002; Hill and Whitham 2014). To our knowledge, no commercially available cultivars of soybean present complete resistance to BPMV infection. Indeed, screening of 52 North American ancestral

soybean lines and examination of all public registration articles through 2002 resulted in no identifiable resistance against BPMV (Wang et al. 2005). Similar results were reported from screening of 303 soybean cultivars grown in the mid-southern region of the United States (Shakiba et al. 2012). Closely related species can show NLR lineage-specific expansions or contractions (Tamborski and Krasileva 2020). Interestingly, our result show that no TNL corresponding to the *I* cluster were present in soybean shedding molecular light on the absence of BPMV resistance in soybean. Common bean and soybean are two closely related species that diverged ~ 19 Mya (Stefanović et al. 2009) and presenting high level of synteny (David et al. 2009; McClean et al. 2010). Consequently, *R-BPMV* from *P. vulgaris* is a precious resource for the development of BPMV resistant soybean. Transferring a resistance gene from one species to another has already proven to be successful in numerous plant species. For example, in *Solanaceae*, susceptible tomatoes to *Xanthomonas* have become resistant after the transfer of *Bs2*, an *R* gene from pepper conferring resistance against the bacteria *Xanthomonas* (Tai et al. 1999). In potato, the *Rpi-vnt1.1* *R* gene from *Solanum venturii* and the *Rpi-mcql* *R* gene from *Solanum mochiquense* were transferred to a potato variety (*Solanum tuberosum*) and showed excellent field resistance against *Phytophthora infestans* (Jones et al. 2014). Concerning monocots, in rice, among 60 *R* genes encoding NLR proteins in maize, sorghum, and *Brachypodium* were transferred into rice lines susceptible to rice blast disease caused by *Magnaporthe oryzae*. Fifteen of them conferred resistance to one or more strains of rice blast disease when expressed in a susceptible rice genotype (Yang et al. 2013). Finally, in the *Fabaceae* family, a CNL from *Cajanus cajan* conferred resistance to *Phakopsora pachyrhizi* when expressed in soybean (Kawashima et al. 2016). Thus, it seems feasible to transfer the *R-BPMV* gene in cultivated soybean genotypes to confer a total resistance against BPMV provided that the downstream signaling pathway is conserved in soybean and that the resistance is efficient against the BPMV strains occurring in US soybean fields. Consequently, identifying which TNL from common bean encodes the *R-BPMV* *R* gene remains an important task for the future. In addition, the resistance mediated by *R-BPMV* is heat-stable up to 35°C, making this *R* gene even more attractive in the context climate change (Meziadi et al. 2021).

Conclusion and future work

The *I* resistance cluster presents many repeated sequences, as illustrated by the difficulty to generate locus specific markers and the poor quality of BAT93 genome assembly, combined with a suppression of recombination. Consequently, these pitfalls prevent the use of a mapped-based cloning strategy to identify the molecular basis of *R-BPMV* gene. Instead, a strategy based

on the identification of mutants would be more relevant as implemented with success for the cloning of the *N* resistance gene in tobacco by insertional mutagenesis with the maize transposon Ac (Whitham et al. 1994). In common bean, a tilling population of BAT93 is available (Porch et al. 2009) and should be screened with BPMV and BCMNV to identify susceptible mutant(s). This will allow to identify which TNL encode(s) *I* and *R-BPMV R* genes and if a single TNL confers resistance to both viruses or if the resistance is conferred by two distinct TNL genes. In parallel, the complete sequence of the *I* cluster should be established in BAT93 and JaloEEP558 taking advantage of recent sequencing technology based on long reads such as PacBio HiFi (Wenger et al. 2019).

Figure captions and Tables

Fig. 1 Fine mapping of the *R-BPMV* gene and identification of flanking markers.

A) Location of *R-BPMV* gene on Linkage Group (LG) B2 based on BAT93 × JaloEEP558 RILs population (Pflieger et al. 2014) and identification of flanking markers IG19M100a, IG19M201b, IG19M401b, and IG19M500b (map distances are indicated between molecular markers and expressed in Kosambi centi Morgans). B) Location on G19833 genome assembly (Schmutz et al 2014) of markers IG19M100a, IG19M201b, IG19M401b, and IG19M500b; physical distances between flanking markers are indicated on the left. C) Chromosome 2 from G19833 genome assembly presenting the distribution of NLR sequences (pink for TIR-NB-LRR, black for CC-NB-LRR). The region corresponding to the *I* TNL cluster is circled in red. The light blue part corresponds to the pericentromeric region.

Fig. 2 Sequence comparison of the *I* cluster region, located on chromosome 2, between *P. vulgaris* G19833 (left) and *P. vulgaris* BAT93 (right).

The sequence from G19833 is based on Schmutz et al. (2014). The sequence from BAT93 is based on Vlasova et al. (2016) (Scaffolds 00018 and 00140) and on the BAC contigs UTG0 and UTG1 (this study). In BAT93, one gap is indicated by two black tipped lines. Putative gene function is listed on the right of each target region in G19833 and BAT93, respectively. Black lines indicate significant homology matches between predicted genes. Physical location of PCR-based markers (IG19M201b, M11RB10, IG19M401b) is indicated by blue lines.

Fig. 3 Microsynteny analysis of the region containing the *R-BPMV* / *I* cluster between *P. vulgaris* G19833 and *Glycine max*. Putative gene function is listed on the right of G19833 target region (middle). In *Glycine max* reference genome target regions, putative gene function is listed on the left and the right of the chromosome 08 and the chromosome 05, respectively. Black lines indicate significant homology matches between predicted genes.

Fig. 4 Phylogenetic analysis of TNLs genes based on amino acid NB domain sequences (from P-loop to the MHD motif).

Colored rectangles correspond to TNLs from *P. vulgaris* G19833 reference genome assembly (PxxGxxxxxx). TNLs from BAT93 BAC contigs (UTG0 and UTG1; this study) are presented in rectangles boxed in green (PvI-xxx). Rectangles boxed in blue correspond to TNLs from *M. truncatula* (Mtxgxxxxxx). TNLs from the *I* cluster are highlighted with a double green arrow. The TNL from *M. truncatula* (Mt8g105820) located in the syntenic region of the *I* cluster is shown by a blue arrow. The TNL from *P. vulgaris* located on chromosome 7 and phylogenetically linked to TNLs from the *I* cluster is shown by a pink arrow.

Table 1 Common bean cultivated genotypes screened for BPMV resistance and for MB29530 marker linked to *R-BPMV* resistance gene.

Gene pool	Cultivated genotype	Phenotype after BPMV-GFP inoculation	Presence/absence of marker MB29530
Mesoamerican	BAT93	R ¹	+
Andean	JaloEEP558	S ²	-
Andean	G19833	S	-
-	Black Valentine	S	-
Andean	Kaboon	S	-
Andean	MDRK	S	-
Andean	Perry Marrow	S	-
Andean	Aiguille verte	S	-
-	Corel	R	+
-	Rosinha	S	-
Mesoamerican	AB136	R	+
Andean	AFN	R	+
Mesoamerican	Cornell 49242	R	+
Mesoamerican	G2333	R	+
Mesoamerican	Michelite	S	-
Mesoamerican	Mz	R	+
Mesoamerican	Pi207262	R	+
Mesoamerican	TO	R	+
Andean	Widusa	R	+

¹ R: no GFP, ² S: GFP in inoculated leaf, + /- : presence or absence of the marker, respectively

Table 2 Common bean wild genotypes screened for BPMV resistance and for MB29530 marker linked to *R-BPMV* resistance gene.

Country	Wild genotype	Phenotype after BPMV-GFP inoculation	Presence/absence of marker MB29530
MEXICO	G11051	S ²	-
ARGENTINA	G19898	S	-
GUETAMALA	G19908	S	-
BOLIVIA	G23442	R ¹	+
ECUADOR	G23582	R	+
COLOMBIA	G24404	R	+
MEXICO	G23429	R	+
HONDURAS	G50722	R	+

¹ R: no GFP, ² S: GFP in inoculated leaf, + /- : presence or absence of the marker, respectively

Acknowledgements We are grateful to Mireille Sévignac and Manon Richard for DNA isolation and RILs genotyping and Sophie Blanchet for her contribution to virus disease scoring. We thank H  l  ne Berg  s and Arnaud Bellec at CNRGV (Toulouse, France) for their help with the BAC library construction. CM was supported by fellowships from the French Research Ministry. JCAD was supported by a fellowship from the Colombian Institute of Education support ICETEX.

Supplementary Information The online version contains supplementary material available at <https://doi.org/xxx>

Authors' contribution statement CM, YCAD and VT developed molecular markers and performed bioinformatic analysis. CM and SP conducted virus resistance evaluation and genetic analysis and were actively involved in the writing of the manuscript. WM constructed and screened the BAC library. AGW, YCAD, SP, ASG and PNM participated in the data analysis, discussion and revised the manuscript. VG designed and supervised the research, discussed the results, coordinated the group work with the help of SP and wrote the manuscript. All authors have read and approved the final manuscript.

Funding This work was supported by the INRAe and IDEEV fundings. IPS2 benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

Data availability Phenotype data generated or analyzed during this study are included in this published article (and its supplementary files). Genotypic data are available from the corresponding author on reasonable request.

Declarations

Conflict of interests On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical standards On behalf of all authors, the corresponding author states that the experiments comply with the current laws of the country in which they were performed.

References

- Altschul S (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
- Ammiraju JSS, Lu F, Sanyal A, et al (2008) Dynamic evolution of oryza genomes is revealed by comparative genomic analysis of a genus-wide vertical data set. *Plant Cell* 20:3191–3209. <https://doi.org/10.1105/tpc.108.063727>
- Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Report* 9:415–415. <https://doi.org/10.1007/BF02672016>
- Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME suite. *Nucleic Acids Res* 43:W39–W49
- Barragan AC, Weigel D (2021) Plant NLR diversity: the known unknowns of pan-NLRomes. *Plant Cell* 33:814–831. <https://doi.org/10.1093/plcell/koaa002>
- Bello MH, Moghaddam SM, Massoudi M, et al (2014) Application of in silico bulked segregant analysis for rapid development of markers linked to Bean common mosaic virusresistance in common bean. *BMC Genomics* 15:903. <https://doi.org/10.1186/1471-2164-15-903>
- Bellucci E, Aguilar OM, Alseekh S, et al (2021) The INCREASE project: Intelligent Collections of food-legume genetic resources for European agrofood systems. *Plant J* 108:646–660
- Bitocchi E, Rau D, Bellucci E, et al (2017) Beans (*Phaseolus* spp.) as a model for understanding crop evolution. *Front. Plant Sci.* 8:722
- Broughton WJ, Hernández G, Blair M, et al (2003) Beans (*Phaseolus* spp.) - Model food legumes. In: *Plant and Soil*. pp 55–128
- Burset M, Guigó R (1996) Evaluation of Gene Structure Prediction Programs. *Genomics* 34:353–367. <https://doi.org/10.1006/geno.1996.0298>
- Cesari S, Bernoux M, Moncuquet P, et al (2014) A novel conserved mechanism for plant NLR protein pairs: the “integrated decoy” hypothesis. *Front Plant Sci* 5:606. <https://doi.org/10.3389/fpls.2014.00606>
- Chalhoub B, Belcram H, Caboche M (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. *Plant Biotechnol J* 2:181–188. <https://doi.org/10.1111/j.1467-7652.2004.00065.x>
- Chen NWG, Sévignac M, Thareau V, et al (2010) Specific resistances against *Pseudomonas syringae* effectors AvrB and AvrRpm1 have evolved differently in common bean (*Phaseolus vulgaris*), soybean (*Glycine max*), and *Arabidopsis thaliana*. *New Phytol* 187:941–956. <https://doi.org/10.1111/j.1469-8137.2010.03337.x>
- Chen NWG, Thareau V, Ribeiro T, et al (2018) Common Bean Subtelomeres Are Hot Spots of Recombination and Favor Resistance Gene Evolution. *Front Plant Sci* 9:1185. <https://doi.org/10.3389/fpls.2018.01185>

699 Choi K, Reinhard C, Serra H, et al (2016) Recombination Rate Heterogeneity within
700 Arabidopsis Disease Resistance Genes. PLOS Genet 12:e1006179.
701 <https://doi.org/10.1371/journal.pgen.1006179>

702 Creusot F, Macadré C, Cana EF, et al (1999) Cloning and molecular characterization of three
703 members of the NBS-LRR subfamily located in the vicinity of the Co-2 locus for
704 anthracnose resistance in Phaseolus vulgaris. Genome 42:254–264.
705 <https://doi.org/10.1139/g98-134>

706 Cronk Q, Ojeda I, Pennington RT (2006) Legume comparative genomics: progress in
707 phylogenetics and phylogenomics. Curr Opin Plant Biol 9:99–103.
708 <https://doi.org/10.1016/j.pbi.2006.01.011>

709 Crute IR, Pink D (1996) Genetics and Utilization of Pathogen Resistance in Plants. Plant Cell
710 8:1747–1755. <https://doi.org/10.1105/tpc.8.10.1747>

711 Dangl JL, Horvath DM, Staskawicz BJ (2013) Pivoting the Plant Immune System from
712 Dissection to Deployment. Science (80-) 341:746–751.
713 <https://doi.org/10.1126/science.1236011>

714 David P, Chen NWG, Pedrosa-Harand A, et al (2009) A Nomadic Subtelomeric Disease
715 Resistance Gene Cluster in Common Bean. Plant Physiol 151:1048–1065.
716 <https://doi.org/10.1104/pp.109.142109>

717 de Haan P, Kormelink R, de Oliveira Resende R, et al (1991) Tomato spotted wilt virus L
718 RNA encodes a putative RNA polymerase. J Gen Virol 72:2207–2216.
719 <https://doi.org/10.1099/0022-1317-72-9-2207>

720 Delgado-Salinas A, Bibler R, Lavin M (2006) Phylogeny of the Genus Phaseolus
721 (Leguminosae): A Recent Diversification in an Ancient Landscape. Syst Bot 31:779–791

722 Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant–pathogen
723 interactions. Nat Rev Genet 11:539–548. <https://doi.org/10.1038/nrg2812>

724 Doyle J., Doyle J. (1991) DNA Isolation from Small Amount of Plant Tissue. Phytochem
725 Bull. <https://doi.org/10.2307/4119796>

726 Drijfhout E (1978) Genetic interaction between Phaseolus vulgaris and bean common mosaic
727 virus with implications for strain identification and breeding for resistance. Versl van
728 Landbouwkond Onderz 1–89

729 Ferreira JJ, Campa A, Pérez-Vega E, et al (2012) Introgression and pyramiding into common
730 bean market class fabada of genes conferring resistance to anthracnose and potyvirus.
731 Theor Appl Genet 124:777–788. <https://doi.org/10.1007/s00122-011-1746-x>

732 Ferrier-Cana E, Geffroy V, Macadré C, et al (2003) Characterization of expressed NBS-LRR
733 resistance gene candidates from common bean. Theor Appl Genet 106:251–261.
734 <https://doi.org/10.1007/s00122-002-1032-z>

735 Ferrier-Cana E, Macadré C, Sévignac M, et al (2005) Distinct post-transcriptional
736 modifications result into seven alternative transcripts of the CC–NBS–LRR gene JA1tr
737 of Phaseolus vulgaris. Theor Appl Genet 110:895–905. <https://doi.org/10.1007/s00122-004-1908-1>
738

- 739 Fisher ML, Kyle MM (1994) Inheritance of resistance to potyviruses in *Phaseolus vulgaris* L.
740 III. Cosegregation of phenotypically similar dominant responses to nine potyviruses.
741 Theor Appl Genet 89–89:818–823. <https://doi.org/10.1007/BF00224503>
- 742 Fisher ML, Kyle MM (1996) Inheritance of resistance to potyviruses in *Phaseolus vulgaris* L.
743 IV. Inheritance, linkage relations, and environmental effects on systemic resistance to
744 four potyviruses. Theor Appl Genet 92:204–212. <https://doi.org/10.1007/BF00223377>
- 745 Fisher RA (1937) The design of experiments. Des Exp
- 746 Flor HH (1956) The Complementary Genic Systems in Flax and Flax Rust. In: Advances in
747 Genetics. pp 29–54
- 748 Foyer CH, Lam H-M, Nguyen HT, et al (2016) Neglecting legumes has compromised human
749 health and sustainable food production. Nat Plants 2:16112.
750 <https://doi.org/10.1038/nplants.2016.112>
- 751 Freyre R, Skroch PW, Geffroy V, et al (1998) Towards an integrated linkage map of common
752 bean. 4. Development of a core linkage map and alignment of RFLP maps. Theor Appl
753 Genet 97:. <https://doi.org/10.1007/s001220050964>
- 754 Garcia T, Duitama J, Zullo SS, et al (2021) Comprehensive genomic resources related to
755 domestication and crop improvement traits in Lima bean. Nat Commun 12:.
756 <https://doi.org/10.1038/s41467-021-20921-1>
- 757 Geffroy V, Creusot F, Falquet J, et al (1998) A family of LRR sequences in the vicinity of the
758 Co-2 locus for anthracnose resistance in *Phaseolus vulgaris* and its potential use in
759 marker-assisted selection. Theor Appl Genet 96:494–502.
760 <https://doi.org/10.1007/s001220050766>
- 761 Geffroy V, Macadré C, David P, et al (2009) Molecular Analysis of a Large Subtelomeric
762 Nucleotide-Binding-Site–Leucine-Rich-Repeat Family in Two Representative Genotypes
763 of the Major Gene Pools of *Phaseolus vulgaris*. Genetics 181:405–419.
764 <https://doi.org/10.1534/genetics.108.093583>
- 765 Geffroy V, Sévignac M, Billant P, et al (2008) Resistance to *Colletotrichum lindemuthianum*
766 in *Phaseolus vulgaris*: a case study for mapping two independent genes. Theor Appl
767 Genet 116:407–415. <https://doi.org/10.1007/s00122-007-0678-y>
- 768 Gepts P, Bliss FA (1985) F1 hybrid weakness in the common bean. J Hered 76:447–450.
769 <https://doi.org/10.1093/oxfordjournals.jhered.a110142>
- 770 Ghabrial SA (1982) Radioimmunosorbent Assay for Detection of Lettuce Mosaic Virus in
771 Lettuce Seed. Plant Dis 66:1037. <https://doi.org/10.1094/PD-66-1037>
- 772 Giesler LJ, Ghabrial SA, Hunt TE, Hill JH (2002) Bean pod mottle virus : A Threat to U.S.
773 Soybean Production. Plant Dis 86:1280–1289.
774 <https://doi.org/10.1094/PDIS.2002.86.12.1280>
- 775 Gonthier L, Bellec A, Blassiau C, et al (2010) Construction and characterization of two BAC
776 libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L.,
777 Asteraceae). BMC Res Notes 3:. <https://doi.org/10.1186/1756-0500-3-225>
- 778 Guindon S, Gascuel O (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large

779 Phylogenies by Maximum Likelihood. *Syst Biol* 52:696–704.
780 <https://doi.org/10.1080/10635150390235520>

781 Hart JP, Griffiths PD (2015) Genotyping-by-Sequencing Enabled Mapping and Marker
782 Development for the By-2 Potyvirus Resistance Allele in Common Bean. *Plant Genome*
783 8:plantgenome2014-09. <https://doi.org/10.3835/plantgenome2014.09.0058>

784 Hill JH, Whitham SA (2014) Control of Virus Diseases in Soybeans. In: *Advances in virus*
785 *research*. United States, pp 355–390

786 Innes RW, Ameline-Torregrosa C, Ashfield T, et al (2008) Differential Accumulation of
787 Retroelements and Diversification of NB-LRR Disease Resistance Genes in Duplicated
788 Regions following Polyploidy in the Ancestor of Soybean. *Plant Physiol* 148:1740–1759.
789 <https://doi.org/10.1104/pp.108.127902>

790 Jones JDG, Witek K, Verweij W, et al (2014) Elevating crop disease resistance with cloned
791 genes. *Philos Trans R Soc B Biol Sci* 369:20130087.
792 <https://doi.org/10.1098/rstb.2013.0087>

793 Katoh K (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment.
794 *Nucleic Acids Res* 33:511–518. <https://doi.org/10.1093/nar/gki198>

795 Kawashima CG, Guimarães GA, Nogueira SR, et al (2016) A pigeonpea gene confers
796 resistance to Asian soybean rust in soybean. *Nat Biotechnol* 34:661–665.
797 <https://doi.org/10.1038/nbt.3554>

798 Kelly JD, Bornowski N (2018) *Biotechnologies of Crop Improvement, Volume 3*. Springer
799 International Publishing, Cham

800 Koinange EMK, Gepts P (1992) Hybrid Weakness in Wild *Phaseolus Vulgaris* L. *J Hered*
801 83:135–139. <https://doi.org/10.1093/oxfordjournals.jhered.a111173>

802 Kosambi DD (1943) The estimation of map distances from recombination values. *Ann Eugen*
803 12:172–175. <https://doi.org/10.1111/j.1469-1809.1943.tb02321.x>

804 Kourelis J, van der Hoorn RAL (2018) Defended to the Nines: 25 Years of Resistance Gene
805 Cloning Identifies Nine Mechanisms for R Protein Function. *Plant Cell* 30:285–299.
806 <https://doi.org/10.1105/tpc.17.00579>

807 Lander ES, Green P, Abrahamson J, et al (1987) MAPMAKER: An interactive computer
808 package for constructing primary genetic linkage maps of experimental and natural
809 populations. *Genomics* 1:174–181. [https://doi.org/10.1016/0888-7543\(87\)90010-3](https://doi.org/10.1016/0888-7543(87)90010-3)

810 Lavin M, Herendeen PS, Wojciechowski MF (2005) Evolutionary Rates Analysis of
811 Leguminosae Implicates a Rapid Diversification of Lineages during the Tertiary. *Syst*
812 *Biol* 54:575–594. <https://doi.org/10.1080/10635150590947131>

813 Lukashin A (1998) GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res*
814 26:1107–1115. <https://doi.org/10.1093/nar/26.4.1107>

815 Märkle H, Saur IML, Stam R (2022) Evolution of resistance (R) gene specificity. *Essays*
816 *Biochem* 66:551–560. <https://doi.org/10.1042/EBC20210077>

817 McClean PE, Mamidi S, McConnell M, et al (2010) Synteny mapping between common bean

818 and soybean reveals extensive blocks of shared loci. *BMC Genomics* 11:184.
819 <https://doi.org/10.1186/1471-2164-11-184>

820 Meyers BC, Dickerman AW, Michelmore RW, et al (1999) Plant disease resistance genes
821 encode members of an ancient and diverse protein family within the nucleotide-binding
822 superfamily. *Plant J* 20:317–332. <https://doi.org/10.1046/j.1365-313X.1999.00606.x>

823 Meyers BC, Kozik A, Griego A, et al (2003) Genome-Wide Analysis of NBS-LRR–Encoding
824 Genes in *Arabidopsis*[W]. *Plant Cell* 15:809–834. <https://doi.org/10.1105/tpc.009308>

825 Meziadi C, Lintz J, Naderpour M, et al (2021) R-BPMV-Mediated Resistance to Bean pod
826 mottle virus in *Phaseolus vulgaris* L. Is Heat-Stable but Elevated Temperatures Boost
827 Viral Infection in Susceptible Genotypes. *Viruses* 13:1239.
828 <https://doi.org/10.3390/v13071239>

829 Meziadi C, Richard MMS, Derquennes A, et al (2016) Development of molecular markers
830 linked to disease resistance genes in common bean based on whole genome sequence.
831 *Plant Sci* 242:351–357. <https://doi.org/10.1016/j.plantsci.2015.09.006>

832 Michelmore RW, Christopoulou M, Caldwell KS (2013) Impacts of Resistance Gene
833 Genetics, Function, and Evolution on a Durable Future. *Annu Rev Phytopathol* 51:291–
834 319. <https://doi.org/10.1146/annurev-phyto-082712-102334>

835 Miklas PN, Kelly JD, Beebe SE, Blair MW (2006) Common bean breeding for resistance
836 against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* 147:105–
837 131. <https://doi.org/10.1007/s10681-006-4600-5>

838 Moghaddam SM, Oladzad A, Koh C, et al (2021) The tepary bean genome provides insight
839 into evolution and domestication under heat stress. *Nat Commun* 12:2638.
840 <https://doi.org/10.1038/s41467-021-22858-x>

841 Morales FJ, Singh SP (1997) Inheritance of the mosaic and necroses reactions induced by
842 bean severe mosaic comoviruses in *Phaseolus vulgaris* L. *Euphytica* 93:223–226.
843 <https://doi.org/https://doi.org/10.1023/A:1002902207395>

844 Mundt CC (2014) Durable resistance: A key to sustainable management of pathogens and
845 pests. *Infect Genet Evol* 27:446–455. <https://doi.org/10.1016/j.meegid.2014.01.011>

846 Ngou BPM, Ding P, Jones JDG (2022) Thirty years of resistance: Zig-zag through the plant
847 immune system. *Plant Cell* 34:1447–1478. <https://doi.org/10.1093/plcell/koac041>

848 Peterson DG, Tomkins JP, Frisch DA, et al (2000) Construction of Plant Bacterial Artificial
849 Chromosome (BAC) Libraries: an Illustrated Guide. *J Agric Genomics* 5:

850 Pflieger S, Blanchet S, Meziadi C, et al (2014) The “one-step” Bean pod mottle virus
851 (BPMV)-derived vector is a functional genomics tool for efficient overexpression of
852 heterologous protein, virus-induced gene silencing and genetic mapping of BPMV R-
853 gene in common bean (*Phaseolus vulgaris* L.). *BMC Plant Biol* 14:.
854 <https://doi.org/10.1186/s12870-014-0232-4>

855 Pflieger S, Blanchet S, Meziadi C, et al (2015) Bean Pod Mottle Virus (BPMV) Viral
856 Inoculation Procedure in Common Bean (*Phaseolus vulgaris* L.). *BIO-PROTOCOL*
857 5:e1524–e1524. <https://doi.org/10.21769/BioProtoc.1524>

858 Pierce WH (1934) Viruses of the bean. *Phytopathology* 24:87–115

859 Porch TG, Blair MW, Lariguet P, et al (2009) Generation of a Mutant Population for
860 TILLING Common Bean Genotype BAT 93. *J Am Soc Hortic Sci J Amer Soc Hort Sci*
861 134:348–355. <https://doi.org/10.21273/JASHS.134.3.348>

862 Richard MMS, Chen NWG, Thareau V, et al (2013) The Subtelomeric khipu Satellite Repeat
863 from *Phaseolus vulgaris*: Lessons Learned from the Genome Analysis of the Andean
864 Genotype G19833. *Front Plant Sci* 4:109. <https://doi.org/10.3389/fpls.2013.00109>

865 Richard MMS, Gratias A, Alvarez Diaz JC, et al (2021) A common bean truncated
866 CRINKLY4 kinase controls gene-for-gene resistance to the fungus *Colletotrichum*
867 *lindemuthianum*. *J Exp Bot* 72:3569–3581.
868 <https://doi.org/https://doi.org/10.1093/jxb/erab082>

869 Richard MMS, Pflieger S, Sévignac M, et al (2014) Fine mapping of Co-x, an anthracnose
870 resistance gene to a highly virulent strain of *Colletotrichum lindemuthianum* in common
871 bean. *Theor Appl Genet* 127:1653–1666. <https://doi.org/10.1007/s00122-014-2328-5>

872 Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist
873 programmers. *Methods Mol Biol* 132:365–386. [https://doi.org/10.1385/1-59259-192-](https://doi.org/10.1385/1-59259-192-2:365)
874 [2:365](https://doi.org/10.1385/1-59259-192-2:365)

875 Rutherford K, Parkhill J, Crook J, et al (2000) Artemis: sequence visualization and
876 annotation. *Bioinformatics* 16:944–945.
877 <https://doi.org/10.1093/bioinformatics/16.10.944>

878 Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid
879 soybean. *Nature* 463:178–183. <https://doi.org/10.1038/nature08670>

880 Schmutz J, McClean PE, Mamidi S, et al (2014) A reference genome for common bean and
881 genome-wide analysis of dual domestications. *Nat Genet* 46:707–713.
882 <https://doi.org/10.1038/ng.3008>

883 Shakiba E, Chen P, Gergerich R, et al (2012) Reactions of Commercial Soybean Cultivars
884 from the Mid South to Soybean Mosaic Virus. *Crop Sci* 52:1990–1997.
885 <https://doi.org/10.2135/cropsci2011.04.0221>

886 Shao Z-Q, Zhang Y-M, Hang Y-Y, et al (2014) Long-Term Evolution of Nucleotide-Binding
887 Site-Leucine-Rich Repeat Genes: Understanding Gained from and beyond the Legume
888 Family. *Plant Physiol* 166:217–234. <https://doi.org/10.1104/pp.114.243626>

889 Shii CT, Mok MC, Temple SR, Mok DWS (1980) Expression of developmental abnormalities
890 in hybrids of *Phaseolus vulgaris* L. *J Hered* 71:218–222.
891 <https://doi.org/10.1093/oxfordjournals.jhered.a109353>

892 Shoemaker RC, Schlueter J, Doyle JJ (2006) Paleopolyploidy and gene duplication in
893 soybean and other legumes. *Curr Opin Plant Biol* 9:104–109.
894 <https://doi.org/10.1016/j.pbi.2006.01.007>

895 Soderlund C, Bomhoff M, Nelson WM (2011) SyMAP v3.4: a turnkey synteny system with
896 application to plant genomes. *Nucleic Acids Res* 39:e68–e68.
897 <https://doi.org/10.1093/nar/gkr123>

898 Stefanović S, Pfeil BE, Palmer JD, Doyle JJ (2009) Relationships Among Phaseoloid
 899 Legumes Based on Sequences from Eight Chloroplast Regions. *Syst Bot* 34:115–128.
 900 <https://doi.org/10.1600/036364409787602221>

901 Steuernagel B, Witek K, Krattinger SG, et al (2020) The NLR-Annotator Tool Enables
 902 Annotation of the Intracellular Immune Receptor Repertoire. *Plant Physiol* 183:468–
 903 482. <https://doi.org/10.1104/pp.19.01273>

904 Tai TH, Dahlbeck D, Clark ET, et al (1999) Expression of the Bs2 pepper gene confers
 905 resistance to bacterial spot disease in tomato. *Proc Natl Acad Sci* 96:14153–14158.
 906 <https://doi.org/10.1073/pnas.96.24.14153>

907 Tamborski J, Krasileva K V. (2020) Evolution of Plant NLRs: From Natural History to
 908 Precise Modifications. *Annu. Rev. Plant Biol.* 71:355–378

909 Thareau V, Dehais P, Serizet C, et al (2003) Automatic design of gene-specific sequence tags
 910 for genome-wide functional studies. *Bioinformatics* 19:2191–2198.
 911 <https://doi.org/10.1093/bioinformatics/btg286>

912 Tock AJ, Fourie D, Walley PG, et al (2017) Genome-Wide Linkage and Association Mapping
 913 of Halo Blight Resistance in Common Bean to Race 6 of the Globally Important
 914 Bacterial Pathogen. *Front Plant Sci* 8:. <https://doi.org/10.3389/fpls.2017.01170>

915 Toro O, Tohme J, Debouck D (1990) Wild bean (*Phaseolus vulgaris* L.): description and
 916 distribution. CIAT

917 Vallejos CE, Astua-Monge G, Jones V, et al (2006) Genetic and Molecular Characterization
 918 of the I Locus of *Phaseolus vulgaris*. *Genetics* 172:1229–1242.
 919 <https://doi.org/10.1534/genetics.105.050815>

920 Vlasova A, Capella-Gutiérrez S, Rendón-Anaya M, et al (2016) Genome and transcriptome
 921 analysis of the Mesoamerican common bean and the role of gene duplications in
 922 establishing tissue and temporal specialization of genes. *Genome Biol* 17:1–18.
 923 <https://doi.org/10.1186/s13059-016-0883-6>

924 Wang Y, Hobbs HA, Hill CB, et al (2005) Evaluation of Ancestral Lines of U.S. Soybean
 925 Cultivars for Resistance to Four Soybean Viruses. *Crop Sci* 45:639–644.
 926 <https://doi.org/10.2135/cropsci2005.0639>

927 Wenger AM, Peluso P, Rowell WJ, et al (2019) Accurate circular consensus long-read
 928 sequencing improves variant detection and assembly of a human genome. *Nat*
 929 *Biotechnol* 37:1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>

930 Whitham S, Dinesh-Kumar SP, Choi D, et al (1994) The product of the tobacco mosaic virus
 931 resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell* 78:1101–1115

932 Wisser RJ, Oppenheim SJ, Ernest EG, et al (2021) Genome assembly of a Mesoamerican
 933 derived variety of lima bean: a foundational cultivar in the Mid-Atlantic USA. *G3*
 934 (Bethesda) 11:. <https://doi.org/10.1093/g3journal/jkab207>

935 Yang S, Li J, Zhang X, et al (2013) Rapidly evolving R genes in diverse grass species confer
 936 resistance to rice blast disease. *Proc Natl Acad Sci* 110:18572–18577.
 937 <https://doi.org/10.1073/pnas.1318211110>

938 Young ND, Debellé F, Oldroyd GED, et al (2011) The Medicago genome provides insight
939 into the evolution of rhizobial symbioses. *Nature* 480:520–524.
940 <https://doi.org/10.1038/nature10625>

941 Zhang J, Li W, Xiang T, et al (2010) Receptor-like Cytoplasmic Kinases Integrate Signaling
942 from Multiple Plant Immune Receptors and Are Targeted by a *Pseudomonas syringae*
943 Effector. *Cell Host Microbe* 7:290–301. <https://doi.org/10.1016/j.chom.2010.03.007>

944

Figures

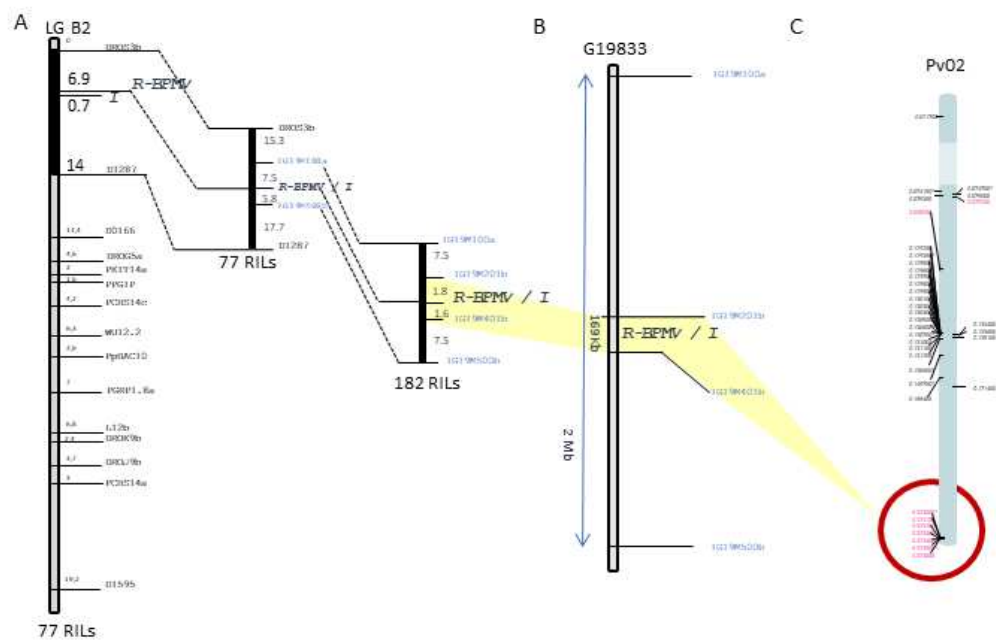


Fig. 1 Fine mapping of the *R-BPMV* gene and identification of flanking markers.
A) Location of *R-BPMV* gene on Linkage Group (LG) B2 based on BAT93 × JaloEEP558 RILs population (Pfllieger et al. 2014) and identification of flanking markers IG19M100a, IG19M201b, IG19M401b, and IG19M500b (map distances are indicated between molecular markers and expressed in Kosambi centimorgans). B) Location on G19833 genome assembly (Schmutz et al 2014) of markers IG19M100a, IG19M201b, IG19M401b, and IG19M500b; physical distances between flanking markers are indicated on the left. C) Chromosome 2 from G19833 genome assembly presenting the distribution of NLR sequences (pink for TIR-NB-LRR, black for CC-NB-LRR). The region corresponding to the *I* TNL cluster is circled in red. The light blue part corresponds to the pericentromeric region.

Figure 1

See image above for figure legend.

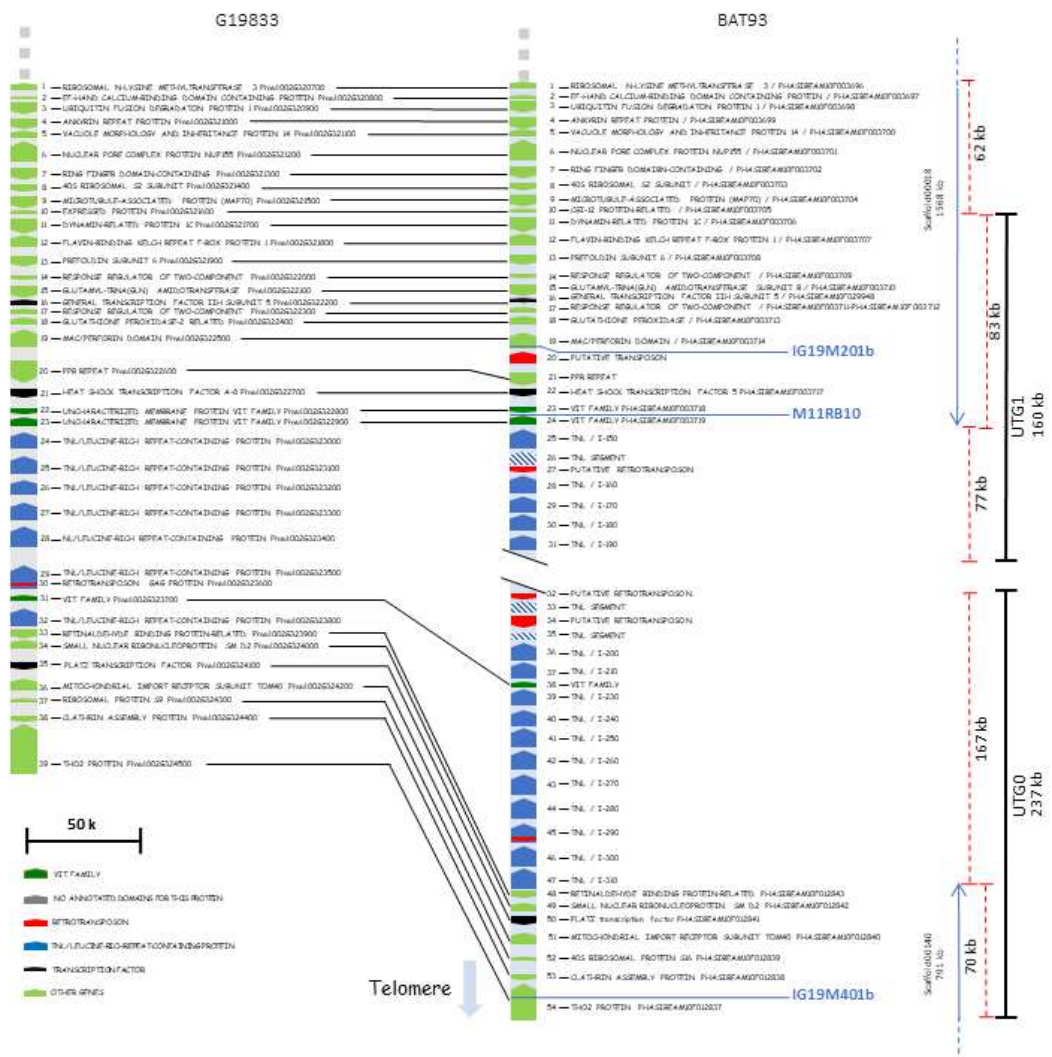


Fig. 2 Sequence comparison of the *I* cluster region, located on chromosome 2, between *P. vulgaris* G19833 (left) and *P. vulgaris* BAT93 (right). The sequence from G19833 is based on Schmutz et al. (2014). The sequence from BAT93 is based on Vlasova et al. (2016) (Scaffolds 00018 and 00140) and on the BAC contigs UTG0 and UTG1 (this study). In BAT93, one gap is indicated by two black tipped lines. Putative gene function is listed on the right of each target region in G19833 and BAT93, respectively. Black lines indicate significant homology matches between predicted genes. Physical location of PCR-based markers (IG19M201b, M11RB10, IG19M401b) is indicated by blue lines.

Figure 2

See image above for figure legend.

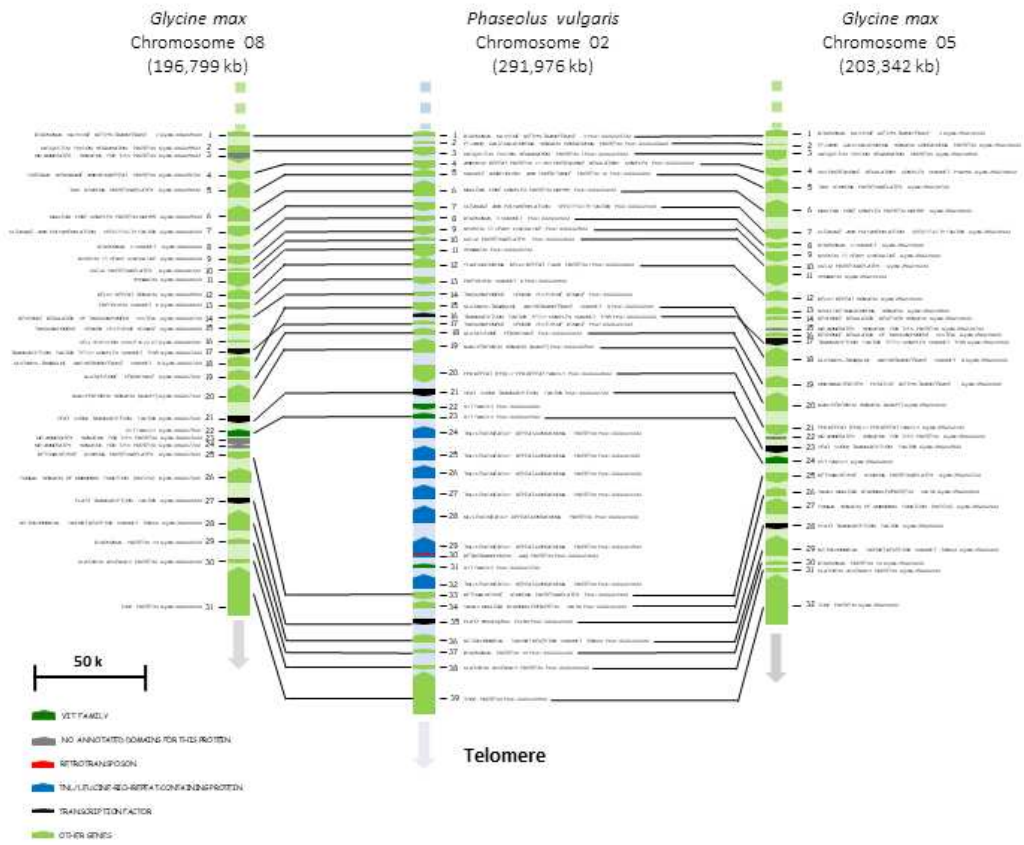


Fig. 3 Microsynteny analysis of the region containing the *R-BPMV* / *I* cluster between *P. vulgaris* G19833 and *Glycine max*. Putative gene function is listed on the right of G19833 target region (middle). In *Glycine max* reference genome target regions, putative gene function is listed on the left and the right of the chromosome 08 and the chromosome 05, respectively. Black lines indicate significant homology matches between predicted genes.

Figure 3

See image above for figure legend.

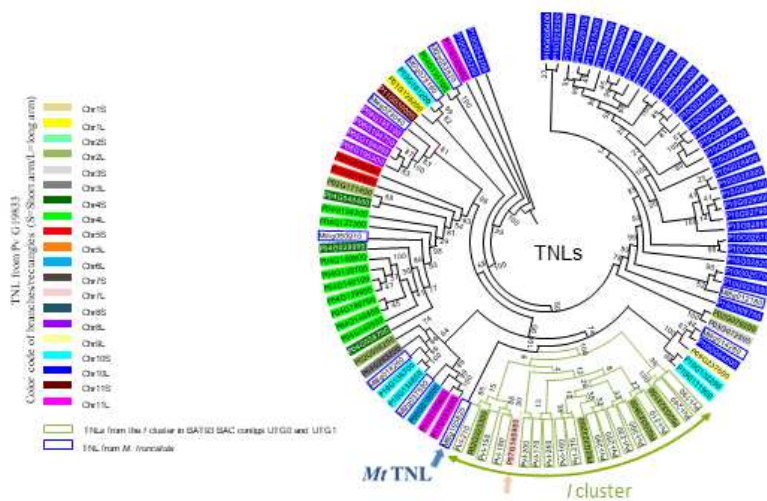


Fig. 4 Phylogenetic analysis of TNLs genes based on amino acid NB domain sequences (from P-loop to the MHD motif).

Colored rectangles correspond to TNLs from *P. vulgaris* G19833 reference genome assembly (PxxGxxxxxx). TNLs from BAT93 BAC contigs (UTG0 and UTG1; this study) are presented in rectangles boxed in green (Pvi-xxx). Rectangles boxed in blue correspond to TNLs from *M. truncatula* (Mtxgxxxxxx). TNL from the *I* cluster are highlighted with a double green arrow. The TNL from *M. truncatula* (Mt8g105820) located in the syntenic region of the *I* cluster is shown by a blue arrow. The TNL from *P. vulgaris* located on chromosome 7 and phylogenetically linked to TNLs from the *I* cluster is shown by a pink arrow.

Figure 4

See image above for figure legend.

Supplementary Files

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

- [SupFigureBACISupfiguresv101402jcaspng.pdf](#)
- [SupTable01032023agwjcaspng.pdf](#)