



# Ma Orthologous Genes in Prunus spp. Shed Light on a Noteworthy NBS-LRR Cluster Conferring Differential Resistance to Root-Knot Nematodes

Cyril van Ghelder, Daniel Esmenjaud, Caroline Callot, Emeric Dubois,  
Marianne Mazier, Henri Duval

## ► To cite this version:

Cyril van Ghelder, Daniel Esmenjaud, Caroline Callot, Emeric Dubois, Marianne Mazier, et al..  
Ma Orthologous Genes in Prunus spp. Shed Light on a Noteworthy NBS-LRR Cluster Conferring Differential Resistance to Root-Knot Nematodes. *Frontiers in Plant Science*, 2018, 9, 10.3389/fpls.2018.01269 . hal-02621840

**HAL Id: hal-02621840**

**<https://hal.inrae.fr/hal-02621840>**

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# *Ma* Orthologous Genes in *Prunus* spp. Shed Light on a Noteworthy NBS-LRR Cluster Conferring Differential Resistance to Root-Knot Nematodes

Cyril Van Ghelder<sup>1\*</sup>, Daniel Esmenjaud<sup>1</sup>, Caroline Callot<sup>2</sup>, Emeric Dubois<sup>3</sup>, Marianne Mazier<sup>4</sup> and Henri Duval<sup>4</sup>

<sup>1</sup> Institut Sophia Agrobiotech, INRA, CNRS, Université Côte d'Azur, Sophia Antipolis, France, <sup>2</sup> Centre National de Ressources Génomiques Végétales, INRA, CNRS, Université de Toulouse, Castanet-Tolosan, France, <sup>3</sup> MGX-Montpellier GenomiX, CNRS, Montpellier, France, <sup>4</sup> Unité de Génétique et Amélioration des Fruits et Légumes (GAFL), INRA, Montfavet, France

## OPEN ACCESS

### Edited by:

Jens Staal,  
Ghent University, Belgium

### Reviewed by:

Daguan Cai,  
Christian-Albrechts-Universität zu Kiel,  
Germany  
Volkan Cevik,  
University of Bath, United Kingdom

### \*Correspondence:

Cyril Van Ghelder  
cyril.van-ghelder@inra.fr

### Specialty section:

This article was submitted to  
Plant Microbe Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 04 May 2018

**Accepted:** 10 August 2018

**Published:** 11 September 2018

### Citation:

Van Ghelder C, Esmenjaud D,  
Callot C, Dubois E, Mazier M and  
Duval H (2018) *Ma* Orthologous  
Genes in *Prunus* spp. Shed Light on  
a Noteworthy NBS-LRR Cluster  
Conferring Differential Resistance  
to Root-Knot Nematodes.  
Front. Plant Sci. 9:1269.  
doi: 10.3389/fpls.2018.01269

Root-knot nematodes (RKNs) are considerable polyphagous pests that severely challenge plants worldwide and especially perennials. The specific genetic resistance of plants mainly relies on the NBS-LRR genes that are pivotal factors for pathogens control. In *Prunus* spp., the *Ma* plum and *RMja* almond genes possess different spectra for resistance to RKNs. While previous works based on the *Ma* gene allowed to clone it and to decipher its peculiar TIR-NBS-LRR (TNL) structure, we only knew that the *RMja* gene mapped on the same chromosome as *Ma*. We carried out a high-resolution mapping using an almond segregating F2 progeny of 1448 seedlings from resistant (R) and susceptible (S) parental accessions, to locate precisely *RMja* on the peach genome, the reference sequence for *Prunus* species. We showed that the *RMja* gene maps in the *Ma* resistance cluster and that the *Ma* ortholog is the best candidate for *RMja*. This co-localization is a crucial step that opens the way to unravel the molecular determinants involved in the resistance to RKNs. Then we sequenced both almond parental NGS genomes and aligned them onto the RKN susceptible reference peach genome. We produced a BAC library of the R parental accession and, from two overlapping BAC clones, we obtained a 336-kb sequence encompassing the *RMja* candidate region. Thus, we could benefit from three *Ma* orthologous regions to investigate their sequence polymorphism, respectively, within plum (complete R spectrum), almond (incomplete R spectrum) and peach (null R spectrum). We showed that the *Ma* TNL cluster has evolved orthologs with a unique conserved structure comprised of five repeated post-LRR (PL) domains, which contain most polymorphism. In addition to support the *Ma* and *RMja* orthologous relationship, our results suggest that the polymorphism contained in the PL sequences might underlie differential resistance interactions with RKNs and an original immune mechanism in woody perennials. Besides, our study illustrates how PL exon duplications and losses shape TNL structure and give rise to atypical PL domain repeats of yet unknown role.

**Keywords:** almond, *Meloidogyne*, nucleotide-binding site - leucine-rich repeat (NBS-LRR), *Prunus*, resistance gene, root-knot nematodes (RKNs), TIR-NBS-LRR (TNL)

## INTRODUCTION

Plant-parasitic nematodes cause huge losses on agriculture worldwide (Nicol et al., 2011). Based on their economic impact, the root-knot nematodes (RKNs) *Meloidogyne* spp. are considered as the number one threat caused by nematodes on crops (Jones et al., 2013). The predominant RKN species are the tropical parthenogenetic mitotic species *Meloidogyne arenaria*, *M. incognita* and *M. javanica* and to a lesser extent the temperate parthenogenetic meiotic *M. hapla*. Nevertheless, minor RKN species such as *M. enterolobii*, *M. chitwoodi*, or *M. fallax* have been recently described as emerging problems for agriculture possibly promoted by climate change or transition in agronomic production (Elling, 2013). The extreme polyphagy of RKNs, characterized by their ability to infect over 5500 plant species, highly increased the interest to identify the actors and generic mechanisms underlying their interaction with plants, especially those involved in plant immunity (Blok et al., 2008).

The plant innate immunity relies on two main processes. The first level involves transmembrane receptors that detect generic molecules usually associated with pathogens. Following this non-specific recognition, this pathway, called pathogen-associated molecular patterns (PAMP) triggered immunity (PTI), activates MAP-kinases cascades leading to plant immunity. This first level of defense can be targeted at different levels by particular pathogen-secreted molecules, known as effectors, leading to inhibition of immunity. In return, effectors can be recognized in a specific manner through a second mechanism so called effector-triggered immunity (ETI) (Jones and Dangl, 2006). The cornerstone of ETI is the NBS-LRR gene family that is involved in plant adaptation to biotic and abiotic stresses. Hundreds of NBS-LRR genes are usually found in genomes of flowering plants (Meyers et al., 1999). They can be subject to deletion through unequal crossing-over or (non)-homologous repair events (Luo et al., 2012) or to expansion involving tandem, ectopic duplications and, to a lesser extent, segmental duplications (Leister, 2004). The dynamic mechanism of gene duplications explained the generic cluster organization of NBS-LRR genes found in plant genomes. This large gene family is split formally into two sub-groups, the TIR-NBS-LRR (TNL) and non-TIR NBS-LRR according to the presence or not of a Toll-interleukin-1 receptor domain in the N-terminal part of the protein (Meyers et al., 1999). Given their modular organization and distinctive domain functions, TNLs act as a multi-function tool for plant immunity. Indeed, the leucine-rich repeat (LRR) domain is polymorphic and often described by its ability to detect pathogens through direct or indirect effector recognition (Jones and Dangl, 2006). The latter process gives rise to some conceptual examples illustrated by the guard protein and decoy models (van der Hoorn and Kamoun, 2008). The central nucleotide-binding site (NBS) domain binds alternatively ADP or ATP leading to conformational modification and by consequence to switch from “OFF” to “ON” states of the protein (Bernoux et al., 2016). The transition between the LRR and the NBS domains is carried by a conserved exon coding for the NBS-LRR Linker (NLL). Once the protein is activated, the TIR domain is released triggering a downstream signaling pathway leading to plant immunity

that is often associated with a programmed cell death called Hypersensitive-like Response (Heath, 2000). However, each domain presumably covers a larger field of competence such as intra or extra-molecular interactions involved in the maintenance of the 3D protein folding, dimerization or alternative effector recognition (Rairdan and Moffett, 2006; Burch-Smith et al., 2007). Besides these three canonical domains, it exists a forth domain, peculiar to TNLs, named post-LRR (PL) domain after its regular position (Van Ghelder and Esmenjaud, 2016). This domain, with unknown role or structure, is usually present in one copy in most of the cloned TNL genes such as *N*, *RPS4*, or *Gro1-4* (Whitham et al., 1994; Gassmann et al., 1999; Paal et al., 2004). Over the last decades, many efforts have been deployed to identify novel genes conferring strong and sustainable natural resistance in plants with agronomical interest. The *Prunus* species are important fruit crops worldwide and, amongst them, *Prunus persica* (peach), *P. cerasifera* (Myrobalan plum) and *P. dulcis* (almond), *qua* perennials, are challenged continuously by RKNs. The common ancestor of the *Prunus* genus emerged approximately 61 Myr in Eastern. The plums (subgenus *Prunophora*) originating from the Eastern Europe and the Middle East, diverged earlier than the almond and peach (subgenus *Amygdalus*) that are native from Western and Eastern Asia (estimated absolute age of 54.6 Myr for *Prunophora* versus 48.9 Myr for *Amygdalus*) (Chin et al., 2014). Although the almond genome is not available yet, the compact peach genome (265 Mbp; Verde et al., 2013) can be considered as the reference genome for *Prunus* species in particular for the diploid ( $2n = 16$ ) species, peach, almond, Japanese and Myrobalan plums, and apricot.

The *Ma* gene, from *P. cerasifera*, was the first TNL gene, conferring resistance to RKNs, cloned in a woody perennial plant (Claverie et al., 2011). Conversely to the *Mi-1* tomato CNL R gene (Milligan et al., 1998), the *Ma* gene confers a heat-stable and complete-spectrum resistance to RKNs (Claverie et al., 2011), and no natural or selected virulent isolates have been detected yet (Khallouk et al., 2013). Besides these attractive biological properties, this gene displays a unique C-terminal structure, which is only found in the *Prunus* and *Malus* genera, made of five PL domains completing the TIR-NBS-LRR canonical part (Van Ghelder and Esmenjaud, 2016). Apart from the *Ma* gene, lying on *Prunus* chromosome 7, several other RKNs *R* genes with more restricted resistance spectra have been identified (Saucet et al., 2016). These are the *RMia* gene from peach, which is most probably a TNL located in a 92-kb interval on the chromosome 2 (Duval et al., 2014) and the *RMja* gene from almond. Interestingly, the latter confers a resistance to *M. javanica* but not to *M. incognita*. A low-resolution mapping strategy carried out on few progenies localized the *RMja* gene on the chromosome 7 (linkage group 7) (Van Ghelder et al., 2010).

The overall goal of this study was to pave the way to identify the molecular determinants involved in the RKN resistance. We have considered three *Prunus* species through accessions that display distinct spectra for resistance to RKNs ranging from complete (plum P.2175/*Ma*) to partial (almond Alnem1/*RMja*) and null (peach Lovell/*Prupe.7G065400*). The specific research objectives were: 1) to localize the *RMja* gene in almond using

high-resolution mapping; 2) to obtain a reliable sequence of the *RMja*-candidate region; 3) to identify the polymorphism that could be involved in the resistance mechanism; 4) and to investigate the evolution of the complete *Ma/RMja* TNL cluster. We showed that the *RMja* gene is located in the *Ma* TNL cluster. Almond NGS genomes, BAC sequences and expression data strongly suggest that the *RMja* gene is the ortholog of the *Ma* plum gene. Additionally, we shed light on the dynamic evolution of the complete *Ma* TNL cluster in plum, peach and almond revealing an atypical focal point for RKN resistance in *Prunus* spp.

## MATERIALS AND METHODS

### Plant and Nematode Materials

The resistant (R) accession *P. dulcis* cv. ‘Alnem1’ controls *M. arenaria* and *M. javanica* but not *M. incognita* and a single gene named *RMja* has been identified to confer resistance against *M. javanica* (Kochba and Spiegel-Roy, 1975; Esmenjaud et al., 2009; Van Ghelder et al., 2010). We produced an intraspecific progeny from almond. Accession ‘Alnem1,’ carrying the homozygous dominant *RMja* gene (*RMja/RMja*), was crossed with the susceptible (S) accession ‘Lauranne’ (recessive for *RMja*; *rmja1/rmja2*) to create a F1 heterozygous progeny. Several Lauranne x Alnem1 F1 hybrids were self-pollinated in greenhouses to produce a segregating F2 progeny (F2 LxA) with three different resistance genotypes (*RMja/RMja*, *RMja/rmja1or2*, and *rmja1or2/rmja1or2*). Shelled almonds were collected and placed onto a culture medium to induce germination.

The population *M. javanica* “Higuera” from Cabrils (Cataluna, Spain) was reared from a single egg mass and maintained on RKN susceptible tomatoes (*Lycopersicon esculentum* cv. ‘St. Pierre’). Its RKN species identity was verified using isoesterase profiling method as defined by Janati et al. (1982).

### Experimental Procedure for Phenotyping

A number of 1448 F2 LxA seedlings, 6-month-old, were planted in pairs in 5-liter pots, placed in greenhouse, irrigated individually every 2 days and grown until harvested for rating at a mean temperature of 25°C (extremes 20–30°C). In parallel, 5-leaf tomato plantlets grown in 250-ml containers were inoculated with 500 24–72-h old juveniles of *M. javanica*. After 2 months, aerial parts of tomato plants were removed and the contaminated soil of one container, including the infected root system, was transplanted into each *Prunus* pot to induce a high- and durable-inoculum pressure. After 6 months, the root system of each *Prunus* plant was carefully examined and rated into two classes: resistant (R) (absence of galls) or susceptible (S) (presence of gall(s)).

### Nucleic Acid Extraction, cDNA Synthesis and Genotyping

The genomic DNA of all F2 plants was extracted using 100 mg of frozen young leaves. Each sample was ground with a mixer

mill in 330 µl of extraction buffer (sorbitol 0.35M, Tris 0.1M, EDTA 5 mM, 4 mg of sodium metabisulfite), 330 µl of lysis buffer (Tris 0.2 M, EDTA 50 mM, NaCl 2 M, CTAB 2%) and 130 µl of sarkosyl 5%. After a chloroform-isoamylalcohol (24:1) procedure, a precipitation with isopropanol and three ethanol washes, the DNA content was eluted in water and treated with RNase.

RNA was extracted using 100 mg of young roots or leaves according to the procedure of Tong et al. (2012) with modifications. Fresh RNA samples were treated with DNase using the Turbo DNA-free kit and evaluated with a Nanodrop. First-strand cDNA were synthesized using the maxima H minus first strand synthesis kit with either specific and oligo (dT) primers. 3 µl of first-strand cDNA were used in PCR using primers defined in **Supplementary Table S1** to conduct transcript detection. The *P. persica* actin gene Prupe.02G235000 was used as control for cDNA synthesis.

PCR were carried out using the MyTaq DNA polymerase kit or the Expand long range kit for longer fragments following the manufacturer’s instructions. SSR primers used (**Supplementary Table S1**) for the genotyping were 5’-fluorescent-labeled with different fluorophores and PCR products were analyzed with a 3130xl Genetic Analyzer.

### Genome Sequencing of the Almond Parents

DNA extracted from leaves of accessions ‘Alnem1’ and ‘Lauranne’ was used to prepare genomic libraries. Tagmentation and PCR amplifications were carried out using the Illumina Nextera DNA sample preparation kit following the supplier’s instructions. The validation of the genomics libraries were performed with a DNA quantification using an Agilent high sensitivity chip and by qPCR. The cluster was carried out inside a flow-cell using the Illumina cluster generation kit. Then, a paired-end 125 sequencing was achieved using an Illumina HiSeq 2500 with a sequence by synthesis (SBS) technique.

Image analysis and base calling were performed using the HiSeq Control Software and Real-Time Analysis component. Demultiplexing was performed using Illumina’s Conversion Software (bcl2fastq 2.17). The quality of the data was assessed using FastQC from the Babraham Institute and the Illumina software SAV (Sequence Analysis Viewer). Potential contaminants were investigated with the FastQ Screen software from the Babraham Institute.

BWA-MEM (v 0.7.12) was used to align reads to the *Prunus persica* genome (v 2.0). MGX-Montpellier GenomiX core facility performed the sequencing of both clones. The raw data of both genomes have been deposited with links to the NCBI BioProject accession number PRJNA448729 (BioSamples numbers SAMN08863554 and SAMN08863555).

### Genomic Library and Bacterial Artificial Chromosome (BAC) Sequencing

High molecular weight (HMW) genomic DNA was prepared from young frozen leaves from accession ‘Alnem1’ as described by Peterson et al. (2000) and Gonthier et al. (2010). Agarose embedded HMW DNA was partially digested with

HindIII, subjected to two size selection steps by pulsed-field electrophoresis using a CHEF Mapper system. DNA was eluted, ligated into the pIndigoBAC-5 HindIII-Cloning Ready vector and transformed in *Escherichia coli* electrocompetent cells. Pulsed-field migration programs electrophoresis buffer and ligation desalting conditions were performed based on Chalhoub et al. (2004). The resulting library represents ~7-fold coverage of accession 'Alnem1' (48 plates, 18,432 BAC clones and mean insert size of 134 kb). BAC clones were spotted on nylon membrane, screened with radioactive labeled probes designed in exons PL1, PL2 and PL5 and revealed by High-density filter reader program. The positive clones were verified by real-time PCR using the specific primers designed in exons PL1, PL2 and PL5. 2 µg of each individual BAC clone of interest were pooled for the construction of a SMRT<sup>®</sup> library using the standard Pacific Biosciences preparation protocol for 10 kb library with PacBio<sup>®</sup> Barcoded Adapters. The pool was then sequenced in one SMRT cell using the P6 polymerase with C4 chemistry. Sequencing was performed on PacBio RS II sequencer. After a demultiplexing step, the sequence assembly was performed following the HGAP PacBio workflow (Chin et al., 2013), and using the SMRT<sup>®</sup> Analysis (v2.3) software suite for HGAP implementation<sup>1</sup>. BAC ends sequences confirmed the position of selected clones, on the peach genome. BAC library construction and screening, and BAC clone sequencing and assembly were performed by the INRA-CNRCV. The data have been deposited with links to NCBI BioProject accession number PRJNA448736.

<sup>1</sup><https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP>

## Sequence Analysis

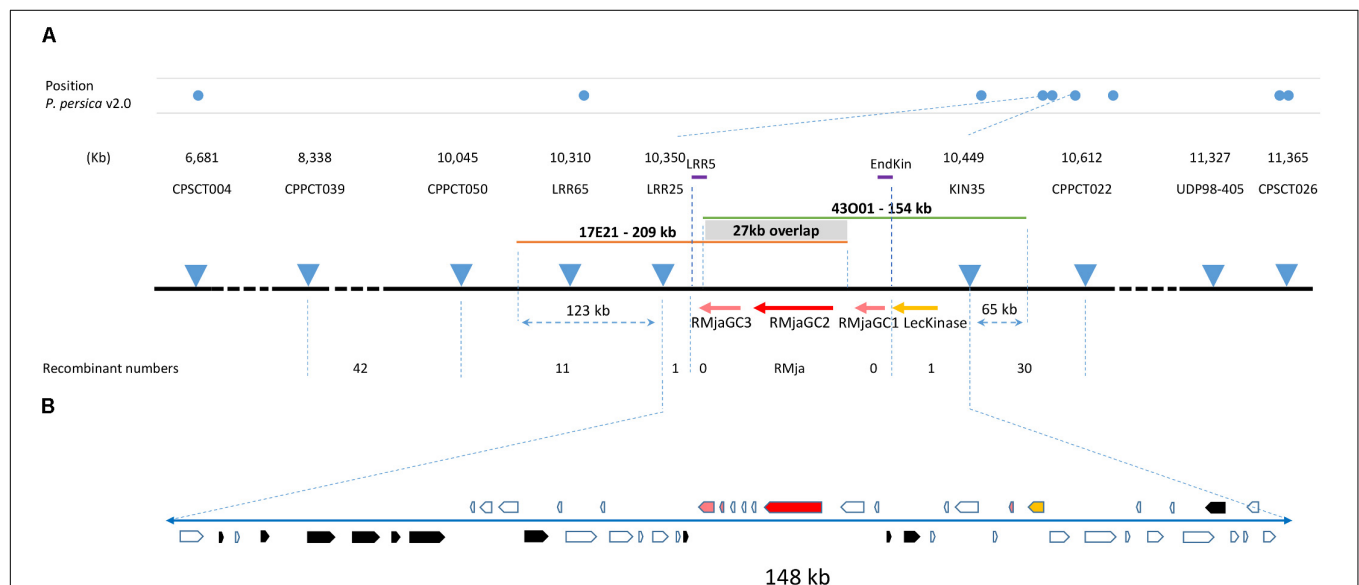
HMM-based gene structure predictions were carried out using FGENESH (Solovyev et al., 2006) and the *P. persica* specific gene-finding parameters. Read alignment and index files (BAM and BAI files), genomic feature files (GFF3) and *P. persica* v2.0.a1 genome<sup>2</sup> were analyzed with the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Alignments, sequence trimmings, mutation rate and phylogenetic analyses were evaluated using MEGA7 (Kumar et al., 2016). Conservative, semi- and non-conservative amino acid substitutions were determined using the Clustal Omega assignment (Sievers et al., 2011). Detection of promoter and regulatory elements were carried out using Cister (Frith et al., 2001) with the *Ma*, *RMjaGC2* and *Prupe.7G065400* genes extended with 7101 bp upstream each of them. We screened 5414, 5414, and 2721 bp upstream the *Ma*, *RMjaGC2* and *Prupe.7G065400* genes, respectively, for W-boxes using PLACE 26.0 (Higo et al., 1999) and PlantTFDB 4.0 (Jin et al., 2017).

## RESULTS

### High-Resolution Mapping of the *RMja* Gene

We localized the *RMja* gene between two SSR markers that cover a window of 0.14 cM. This region corresponds to a physical interval of 99 kb in the peach genome v2.0.a1. The high-resolution mapping strategy conducted on 1448 total F2

<sup>2</sup>[ftp://ftp.bioinfo.wsu.edu/species/Prunus\\_persica/Prunus\\_persica-genome.v2.0.a1/assembly/](ftp://ftp.bioinfo.wsu.edu/species/Prunus_persica/Prunus_persica-genome.v2.0.a1/assembly/)



**FIGURE 1 |** High-resolution mapping of the *RMja* region. **(A)** Molecular markers used in this study and their position in the peach genome v2.0 are indicated at scale with blue dots. The numbers of recombinants are indicated below the respective markers and the positions of the two BACs, 17E21 and 43001, are indicated with orange and green lines, respectively. **(B)** Magnification of the final 148-kb interval in almond containing the *RMja* gene. The *RMjaGC2* gene (*Ma* ortholog) is highlighted in red, the *RMjaGC1* and *RMjaGC3* genes in pink, the lectin tyrosine kinase in yellow and transposable elements are indicated in black. The white arrows indicate predicted peptides in the interval that are described in **Supplementary Table S3**.

almond seedlings revealed a single recombination event for each of the flanking markers, LRR25 and KIN35. Internal primers named LRR5 and EndKin designed within this final interval, amplified fragments that eliminated any recombination events allowing us to refine the *RMja* gene localization (**Figure 1A**). This mapping was actually performed in two steps. In a first step, we used a progeny composed of 907 F2 seedlings that clearly split into 625 R individuals and 282 S individuals. We tested the S and R parental accessions together with twenty hybrids to select reliable and polymorphic 5'-fluorescent-labeled SSR markers spread in this LG7 region from the peach genome (Verde et al., 2013). Genotyping the 907 progenies with the flanking markers CPPCT039 and CPPCT022 (**Supplementary Table S1**) revealed 27 and 16 recombinant individuals on each side of the gene, respectively. In a second step, the 541 F2 individuals obtained segregated into 411 R: 130 S individuals (3 R: 1S ratio;  $\chi^2 = 0.27$ ;  $P = 0.6$ ). The overall individuals segregated into 1036 R: 412 S individuals ( $\chi^2 = 9.2$ ;  $P = 0.002$ ). In the candidate interval, we developed three new SSR markers, LRR65, LRR25 and Kin35 (**Supplementary Table S1**) that revealed eleven, one and one recombinant individuals in the total progeny, respectively (**Figure 1A**).

## The *RMja* Gene Is Located in a Region Containing the *Ma* Ortholog

In the peach genome, a further examination of the *RMja* final interval marked out by the LRR25 and KIN35 markers, revealed eleven candidate genes named *Prupe.7G065100* to *Prupe.7G066100* (**Supplementary Table S2**). Among these candidate genes, a functional annotation identified one tyrosine kinase (*Prupe.7G065700*) and four TNLs (*Prupe.7G065300* to *Prupe.7G065600*) including the *Ma* ortholog (*Prupe.7G065400*). We sequenced the genome of the parental accessions 'Alnem1' (R) and 'Lauranne' (S) to further characterize the candidate region in almond. The phyletic proximity between peach and almond allowed us to map accurately the reads on the *P. persica* genome assembly (v2.0.a1). All over the region, the alignment of both the R and S almond reads onto the *P. persica* genome indicated a similar pattern involving alternation between syntenic regions with peach, and indels of unknown size (**Figure 2A**). We identified two profiles in the syntenic regions showing either (i) a standard coverage ( $20\times - 50\times$ ) and moderate nucleotide polymorphism or (ii) a high coverage ( $>100\times$ ) and nucleotide polymorphism. The first pattern corresponds to orthologous loci whereas the second, presumably, reflects the misalignment of ectopic transposable elements (TEs). Magnification of the region spanning from the first TNL (*Prupe.7G065300*) to the lectin kinase (*Prupe.7G065700*), arbitrarily named R region, highlights the absence of the ortholog of the TNL *Prupe.7G065500* in almond (**Figure 2B**). The almond orthologs of *Prupe.7G065700* (lectin kinase) and *Prupe.7G065400* (*Ma* ortholog TNL) display a significant level of homology (illustrated in gray in **Figure 2B**) unlike the orthologs of *Prupe.7G065300* and *Prupe.7G065600* showing nucleotide polymorphism and numerous indels. More specifically, the *Prupe.7G065400* gene shows conserved structural features and low polymorphism in predicted coding regions in

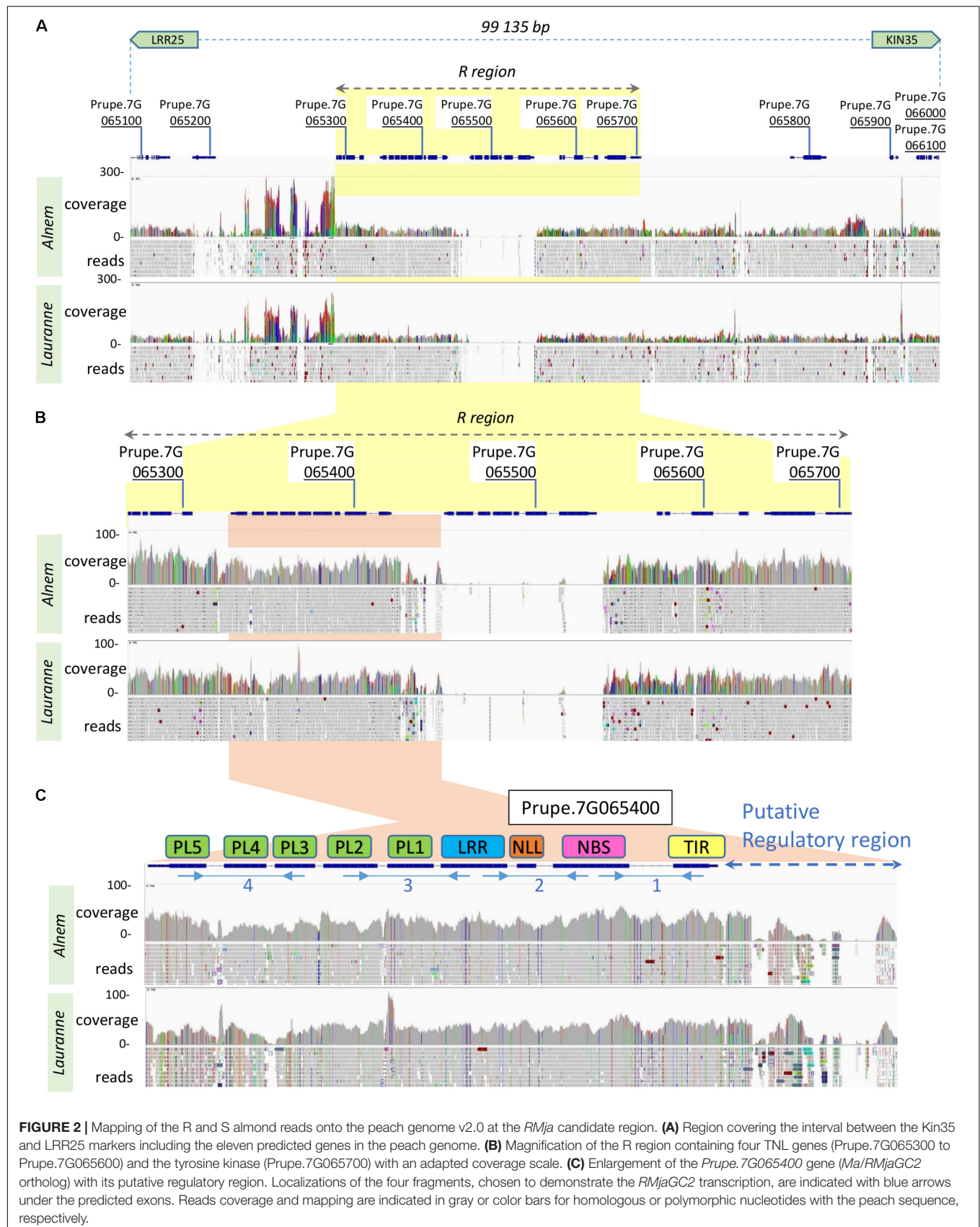
either the homozygote 'Alnem1' (R), or 'Lauranne' (S), or peach 'Lovell' accessions. Only the upstream putative regulatory region displays substantial divergences between both almonds and the peach accessions (**Figure 2C**).

## BAC Sequences and Expression Data Strongly Suggest That *RMja* Is the Ortholog of the *Ma* Gene

We selected BAC clones that correspond to the region of interest from a genomic library constructed from the R accession 'Alnem1.' We sequenced two overlapping BAC clones ('17E21' and '43O01') [PRJNA448736] covering the KIN35-LRR25 interval, using long-read sequencing technology to overcome potential misassembly due to repeated sequences. These two sequences meet the quality criteria with a quality value  $QV > 48$  meaning a high confidence in nucleotide identification. The BAC clone '43O01' contains an insert of 154 kb that can be aligned from 65 kb upstream the KIN35 marker to the *RMjaGC3* gene. The second BAC clone sequence '17E21' is 209 kb long and maps from the *RMjaGC1*-*RMjaGC2* intergenic region to 123 kb downstream the LRR25 marker (**Figure 1A**). The overlapping region between the two BAC clones covers a sequence of 27 kb that displays a similarity of 99.97 %. Only eight nucleotide deletions were found and localized in five SSR regions in the 27 kb, whereas the remainder 190 kb identified SSRs were strictly identical. The fragment length between the KIN35 and LRR25 markers is 49 kb longer in almond (**Supplementary Table S5**) than in peach.

Gene finding analyses and functional annotations revealed 48 predicted ORFs including one lectin kinase, only three TNLs of different sizes, small peptides (potential artefacts) and TEs that explain the size variation in this region (**Figure 1B** and **Supplementary Table S3**). The lectin kinase that precedes the TNL cluster is highly conserved between plum, peach and almond ( $>98\%$  amino acid identity). The protein is composed of a signal peptide, an extra-cellular carbohydrate-binding domain, a transmembrane portion and an intra-cellular tyrosine kinase domain (**Supplementary Table S4**). This gene is the ortholog (54.8% identity) of the *Arabidopsis thaliana* gene coding protein At2g41890 assigned to the GO biological process of pollen recognition and self-incompatibility. A blast analysis revealed that the *Ma* and lectin kinase loci existed separately in Rosales species but only co-localized in *Prunoideae* species as there are only found physically close to each other in *Malus*, *Pyrus* and *Prunus* genomes (**Supplementary Table S6**).

Immediately after the kinase, we found a TIR domain-only protein (*RMjaGC1*). The analysis of the immediate downstream sequence revealed a disrupted NBS domain followed by parts of retrotransposons suggesting a progressive gene loss. Except the *RMjaGC2* gene, the candidate interval then only contains a last TNL, *RMjaGC3*. This gene is predicted to be split into a TIR domain-only protein followed by NBS-LRR protein in which, the LRR domain is truncated (232 aa) due to a frameshift mutation inducing a premature stop codon. Furthermore, a retrotransposon fragment, present on the minus strand, disrupts its putative C-terminal part (**Figure 1B**).



The *RMjaGC2* gene is the sole TNL in the R candidate-region to display a complete structure with full-length domains and to be constitutively transcribed. Indeed, we carried out transcript analysis in the 'Alnem1' R and 'Lauranne' S parental accessions and in several R F2 individuals. In Alnem1 as in R F2 individuals, we successfully amplified and sequenced four fragments from the *RMjaGC2* transcript, using primers designed across exons (**Supplementary Table S1** and **Figure 2C**) and cDNA from leaves and roots as illustrated in **Figure 3**. We also detected the *RMjaGC2* transcript in the single recombinant between KIN35 and EndKin markers, which was phenotypically resistant. Unlike R accessions, the same manipulation failed to detect any *RMjaGC2* transcripts in 'Lauranne' (S). In parallel, we failed to amplify any transcripts of *RMjaGC1* and *RMjaGC3* genes in 'Alnem1.' A similar approach detected the transcripts of the *Ma* and *Prupe.7G065400* genes in leaves of RKN-free plants of each of the accessions *P. cerasifera* cv. 'P.2175' (the *Ma* donor) and *P. persica* cv. 'Lovell,' respectively, suggesting a constitutive expression and an absence of tissue-specificity in the three genes (**Figure 3**). Moreover, a blast analysis using the *Ma*, *RMjaGC2* and *Prupe.7G065400* transcript sequences as queries against RNAseq and ESTs data from diverse studies, confirmed the transcription of this gene in (i) roots and leaves of *P. persica* and (ii) in various cultivars from different *Prunus* species (**Supplementary Table S6**).

An *in silico* analysis predicted a regulatory region approximately 1kb upstream the start codon of the *RMjaGC2* gene while no consistent *cis*-element binding sites were detected in the vicinity of the *RMjaGC1* and *RMjaGC3* genes. The predicted *RMjaGC2* regulatory region appeared to be very

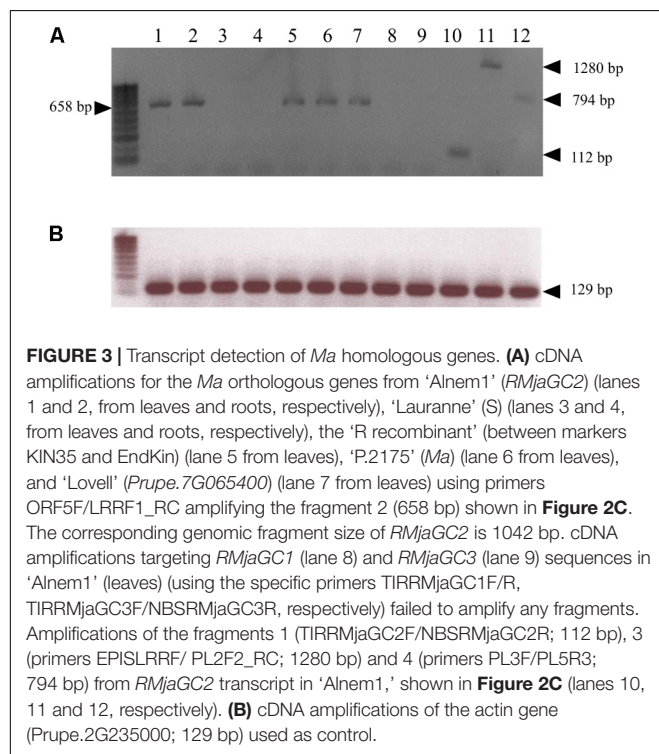
similar to that of peach but quite different from that of plum where multiple regulatory regions are detected (**Supplementary Figure S1**). In parallel, we screened the upstream region of the *Ma*, *RMjaGC2* and *Prupe.7G065400* genes for putative W-boxes (WRKY binding site). Two independent predictions identified an enrichment (>40%) of putative W-boxes localized upstream the *Ma* and *RMjaGC2* genes compared to the single one present upstream the *Prupe.7G065400* (**Supplementary Figure S1**). In peach, an insertion of a TNL gene (*Prupe.7G65500*) has probably replaced a part of the W-boxes in this region.

## The PL Domains Contain Substantial Variations Between *RMjaGC2* and Its Orthologs

To date, the *Ma* gene is the sole gene conferring a full resistance to all RKN species tested including *M. incognita*, *M. arenaria* and *M. javanica*. It belongs to a cluster of three TNL genes (of which two display structural defects) preceded by a tyrosine kinase. The three orthologous genes (*Ma*, *RMjaGC2*, *Prupe.7G065400*) are highly similar (>96% CDS identity). While the *Ma* coding sequence displays a similar identity with peach and almond, its predicted protein shows a slightly higher identity to *RMjaGC2* (94.41%) than to *Prupe.7G065400* (94.12%) (**Table 1**).

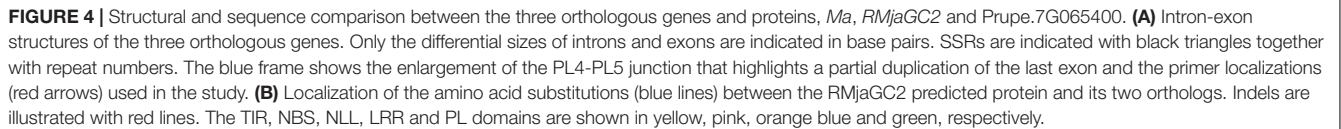
A gene structure analysis revealed a disparity localized in the first and last intronic regions of these genes. In *Ma*, the first intron between the TIR and NBS-coding exons displays an approximate 1.5 kb fragment that is absent in both *RMjaGC2* and *Prupe.7G065400* (**Figure 4A**). Each PL-coding exon is preceded by dinucleotide SSRs implying polymorphic areas. The final intron displays a differential pattern. Indeed, we noticed that in the *RMjaGC2* and *Prupe.7G065400* genes, the final intron exhibits a duplication of a fragment containing a part of the intron and 32 nucleotides of the last exon including the splicing acceptor site (**Figure 4A**). This singularity might induce alternative splicing and/or truncated mRNA. Using selective primers, PCR experiments carried out with Alnem1 cDNA, failed to amplify any alternative transcripts and only the predicted PL4-PL5 was detected (**Figure 4A**). Nevertheless, we cannot exclude that alternative forms might exist in low numbers or might be induced under specific conditions.

At the peptide level, we identified 110 amino acid mutations that are specific to *Ma*, 12 that are specific to *RMjaGC2* and 14 that are only present in *Prupe.7G065400*. Unlike the polymorphism detected between *RMjaGC2* and



**TABLE 1 |** Percent identity matrix between the three orthologous genes, CDS and proteins, *Ma*, *Prupe.7G065400* and *RMjaGC2*.

|                |                       | <i>Prupe.7G065400</i> | <i>Ma</i> |
|----------------|-----------------------|-----------------------|-----------|
| Gene UTR incl. | <i>RMjaGC2</i>        | 98.42                 | 91.41     |
|                | <i>Prupe.7G065400</i> |                       | 91.52     |
| CDS            | <i>RMjaGC2</i>        | 99.20                 | 96.65     |
|                | <i>Prupe.7G065400</i> |                       | 96.67     |
| Protein        | <i>RMjaGC2</i>        | 98.78                 | 94.41     |
|                | <i>Prupe.7G065400</i> |                       | 94.12     |



We established a repertoire of mutations that are (1) only present in Ma (Ma specific) compared to both RMjaGC2 and Prupe.7G065400, (2) specific to RMjaGC2 (RMjaGC2 specific) or (3) common to both Ma and RMjaGC2 and so differing from Prupe.7G065400 (Ma-RMjaGC2 specific). This latter category might highlight crucial amino acid for resistance whereas the first two categories might be associated to spectrum determinants of the resistance. Each mutation is assigned to conservative, semi-conservative or non-conservative mutation subsets as such characteristics often have a differential impact on protein functioning (**Table 2**). In Ma-RMjaGC2 specific mutations, we detected two non-conservative events in both PL1 (T1103R; G1150R) and PL2 (P1408Q; T1423R) domains that may influence the resistance mechanism. The non-conservative mutations that are specific to Ma are mostly localized in the LRR, PL1, PL2, PL3 and PL4 domains with a maximum of 12 non-conservative amino acid mutations in the PL2. RMjaGC2 displays an alanine deletion in the PL1 and one non-conservative mutation in PL2, PL3 and PL4 (S1281L, Q1570P and T1668L, respectively) that may be in relation to its selective spectrum of resistance.

## Ma Cluster Evolution

The generation of the almond sequence, together with its syntenic regions in plum and in peach, offered us the opportunity to study the evolution of a cluster containing atypical and functional TNL genes. The pairwise analysis of the ten different TNL genes detected in the cluster revealed two conserved loci over the three species (**Figure 6**). The first conserved locus includes the *MaGC3*, *Prupe.7G065300* and *RMjaGC3* genes that display similar structure with intact TIR, NBS and NLL domains but a truncated LRR domain and no PL domains due to TEs insertion or frameshift mutations

TABLE 2 | Distribution of the conservative (C), semi-conservative (S) and non-conservative (N) amino acid mutations over the different domains.

| Mutations             | Total |    |    | TIR |   |   | NBS |   |   | NLL |   |   | LRR |   |   | PL1 |   |   | PL2 |   |    | PL3 |   |   | PL4 |   |   | PL5 |   |   |
|-----------------------|-------|----|----|-----|---|---|-----|---|---|-----|---|---|-----|---|---|-----|---|---|-----|---|----|-----|---|---|-----|---|---|-----|---|---|
|                       |       |    |    |     |   |   |     |   |   |     |   |   |     |   |   |     |   |   |     |   |    |     |   |   |     |   |   |     |   |   |
|                       | C     | S  | N  | C   | S | N | C   | S | N | C   | S | N | C   | S | N | C   | S | N | C   | S | N  | C   | S | N | C   | S | N | C   | S | N |
| Ma specific           | 53    | 13 | 44 | 3   | 0 | 3 | 3   | 0 | 0 | 4   | 1 | 0 | 5   | 1 | 6 | 8   | 7 | 7 | 12  | 1 | 12 | 8   | 2 | 7 | 6   | 1 | 9 | 4   | 0 | 1 |
| RMjaGC2 specific      | 2     | 5  | 5  | 0   | 0 | 1 | 1   | 0 | 0 | 0   | 1 | 0 | 1   | 2 | 0 | 0   | 1 | 1 | 0   | 0 | 1  | 0   | 0 | 1 | 0   | 0 | 1 | 0   | 1 | 0 |
| RMjaGC2 & Ma specific | 8     | 2  | 4  | 2   | 0 | 0 | 1   | 1 | 0 | 0   | 0 | 0 | 2   | 0 | 0 | 1   | 0 | 2 | 0   | 1 | 2  | 0   | 0 | 0 | 0   | 0 | 0 | 2   | 0 | 0 |

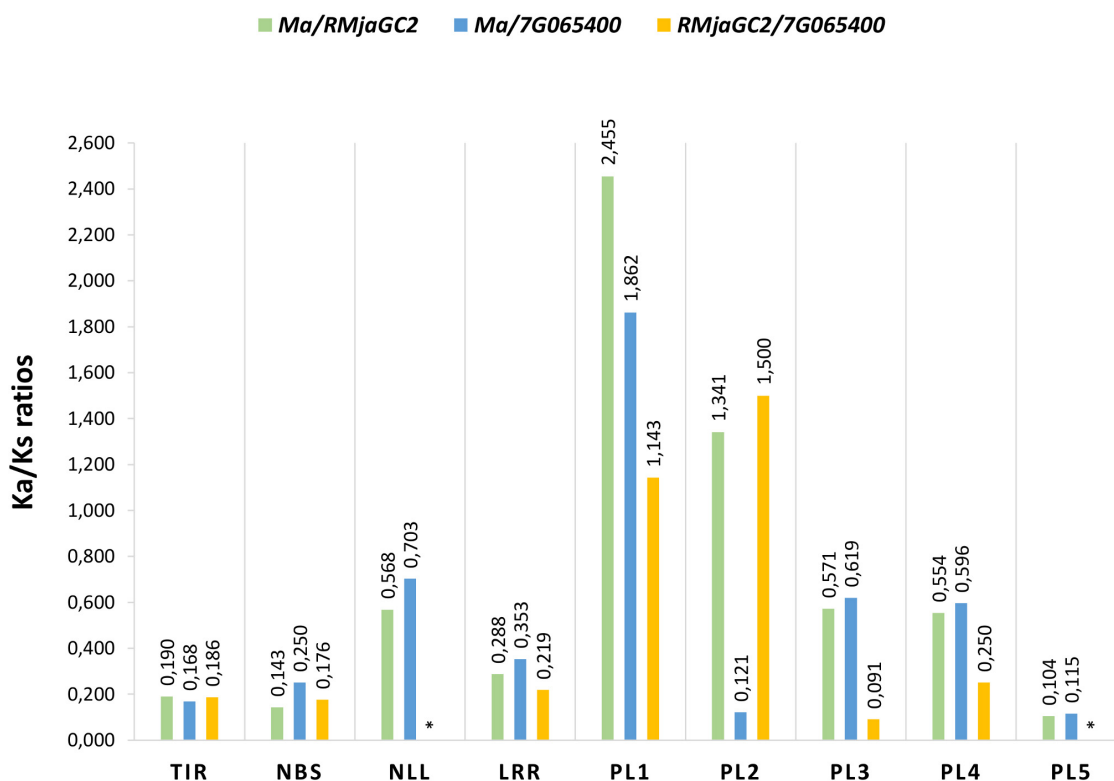
inducing premature stop codons. The second conserved locus includes the *Ma*, *RMjaGC2* and *Prupe.7G065400* genes, characterized by a complex structure. The *Prupe.7G065500*, *Prupe.7G065600* and *RMjaGC1* genes are related to each other and underwent severe reorganizations including TEs insertions or duplications leading to chimeric structures. They emerged either from the *Prunus* ancestor and have been lost in plum, or they are posterior to the split between plums and *Amygdalus*.

In plum, pairwise and phylogenetic analyses of the diverse domains revealed that the *MaGC2* gene recently emerged from partial duplications of the *Ma* and *MaGC3* genes with internal rearrangements. Indeed, the reorganization involves a TIR domain originating from *MaGC3*, and a remaining sequence originating from the *Ma* gene except two lost exons (PL3 and PL4). In *RMjaGC2* and its orthologs, the PL1 domain is the closest PL domain to the outgroup (the PL domain of the tobacco *N* gene). It probably duplicated to give birth to PL2, which then diverged. PL3, PL4 and PL5 domains display a significant identity between each other and are closer to PL1 than the PL2 domains (Figure 6 and Supplementary Figure S2). These results suggest that the PL1 domain underwent exon duplication to generate (i) PL2, a long time ago, and (ii) one of PL3, PL4 or PL5, which rapidly duplicated itself to produce this unique structure.

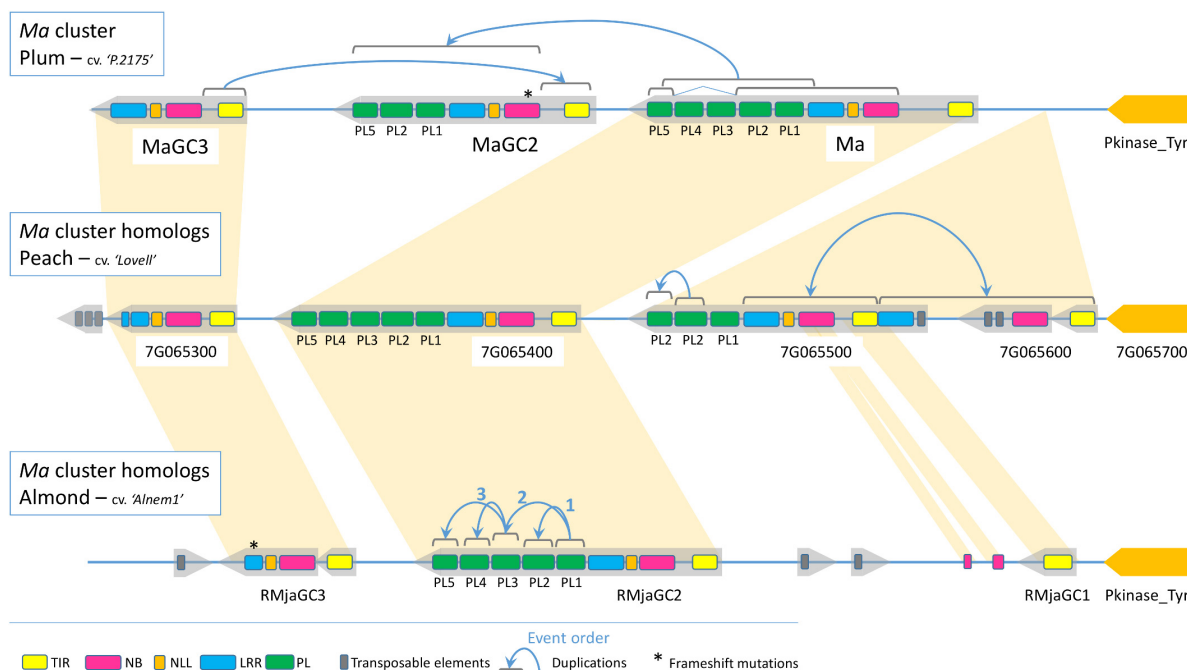
DISCUSSION

The *RMja* Gene Is Presumably the *Ma* Ortholog

Our high-resolution mapping together with the peach genome and *Ma* gene resources helped us to precisely locate the *RMja* gene. This approach, based on a large progeny, points out a limited window in which very few predicted ORFs are credible candidate genes to confer such a resistance. In ‘Alnem1,’ the precise annotation of the *RMjaGC2* gene and its environment together with expression data strongly suggest that the *RMjaGC2* gene is *RMja* and so confers resistance to *M. javanica*. Nevertheless, the poor *Prunus* ability to transformation, the peculiar recalcitrance of almond to produce roots spontaneously together with the large size of *RMjaGC2* are so many obstacles to engineer *RMja*-transformed roots and so validate the *RMjaGC2* gene. Its transfer in species with a more favorable genetic background would be an alternative. In this perspective, preliminary works to transfer *Ma* in tomato failed to provide any resistance, even though a constitutive transcription has been detected (H. Duval, unpublished). Missing components in *Solanaceae* probably explains the unsuccessful transfer of resistance, which seriously restrains this approach at least in these models. The dual transfer of *Arabidopsis* NBS-LRR genes, *RRS1* and *RPS4*, circumvented the restricted taxonomic functionality (Tai et al., 1999) in *Solanaceae* species (Narusaka et al., 2013). In the future, the identification of a *Ma/**RMja* partner could help consequently to break down the technical barriers that are obstructing deeper functional analyses.



**FIGURE 5 |** Non-synonymous (Ka) to synonymous (Ks) substitution ratios (Ka/Ks) in the different domains between the three orthologs *Ma*, *RMjaGC2* and *Prupe.7G065400*. \*Indicates absence of Ka/Ks ratio due to a null Ks.



**FIGURE 6 |** Schematic representation of the R region containing the *Ma* TNL cluster. Structural composition and rearrangements occurring in the ten TNL homologous genes in peach, plum and almond.

## Insight Into the Molecular Determinants Involved in RKN Resistance

The *Ma* gene is not circumvented so far and confers resistance to mitotic species *M. incognita*, *M. javanica*, *M. arenaria*, *M. ethiopica*, *M. enterolobii* and the meiotic *M. floridensis* (Claverie et al., 2011). A recent study showed that *Ma* and *RMja* resistance spectra are closer than expected. Indeed, the sole species uncontrolled by *RMja* (*M. incognita* and *M. floridensis*) (D. Esmenjaud, unpublished) are also close in the phylogenetic tree of the *Meloidogynidae* family (Holterman et al., 2009). The weak polymorphism between *RMjaGC2* and its peach ortholog suggests that few mutations may confer resistance to most RKNs. The differential specificity between *Ma* and *RMja* may be due either to the gain of recognition of a ubiquitous effector from *Meloidogyne* spp. in *Ma* or to the loss of recognition of an effector specific to *M. incognita* and *M. floridensis* in *RMja*. In the case of a direct recognition of effectors by NBS-LRRs, few mutations may lead to extended recognition specificity or, on the contrary, to a failure of recognition due to a rapid co-evolutionary arms race (Karasov et al., 2014). In this context, we should observe among *Prunus* accessions a panel of alleles conferring variable resistance interactions to RKNs for species specificity and/or resistance strength.

Structurally, the three genes *RMjaGC2*, *Ma* and *Prupe.7G065400* display an identical repetition of five PL domains. The cost for the plant to produce and conserve such massive C-terminal ends in *Prunus* species is in line with an essential role of these repeats. PLs may function as baits mimicking more or less fitted RKN effector targets and thus may reveal quite original immune mechanisms in woody perennial plants. Screening identified RKN effectors [reviewed with their potential targets in Truong et al. (2015)] for *Ma* recognition is a strategy that should help to clarify the RKN resistance mechanism in *Prunus*. Even though the role of the PL domain is still unclear, mutation in a conserved PL-domain motif (i.e., the second motif defined in Van Ghelder and Esmenjaud, 2016) of the Arabidopsis *RPS4* gene, impaired the dual-*RPS4/RRS1* mediated resistance suggesting a role of this domain in immune signaling and/or partner interactions (Sohn et al., 2014).

## Position and Role of the 5-PL-Containing Genes in the Resistance Mechanism

*Ma* and *RMja* are single homologous genes conferring resistance to different RKN species. Nevertheless, other partners could also be actively involved in this process. The orthologous genes *Rpi-blb2* and *Mi-1.2* from potato and tomato, respectively, conferred resistance to certain RKNs, aphids and white flies (van der Vossen et al., 2005). Recent studies showed that these two genes required other NBS-LRRs (NLR required for cell death-helpers or NRC-helpers) to function. Besides, transfers of diverse tomato NRC-helpers together with *Rpi-blb2/Mi* in NRC-silenced *N. benthamiana* showed contrasting rescues of *Rpi-blb2/Mi*-mediated cell death (Wu et al., 2017). This highlights the influence of the NBS-LRR sensor-helper affinity to transfer

successfully a complete resistance and interrogates about a model in which *Ma*, *RMja* or *Prupe.7G065400* sensors would confer distinct resistance spectra in accordance with a divergent affinity between sensors and helper(s) in plum, almond, and peach.

As crucial elements of PTI, kinase proteins are effector targets, which will modify and alter kinases activity favoring pathogen growth and development. The conserved lectin S-domain receptor kinase in the R region, which mostly segregates with *Ma* orthologs, is classified within the *Galanthus nivalis* agglutinin (GNA) homologs (Eggermont et al., 2017). Some members are known to be upregulated in presence of pathogens while others are presumably involved in calcium influx, which is one of the PTI characteristics. At the same time, many targets of RKN effectors are unknown, even in the well-described *M. incognita* calreticulin effector Mi-CRT, which is crucial for infection by possibly altering calcium homeostasis (Jaouannet et al., 2013). In this perspective, the lectin kinase might be targeted by RKN effectors and monitored by *Ma/RMja* as illustrated in the kinase/NBS-LRR (PBS1/ RPS5; Pto/Prf) models in which kinase modifications induced by bacterial effectors are detected by NBS-LRRs inducing immune responses (Zhang et al., 2010; Ntoukakis et al., 2014). Interestingly, the CNL *Prf* also lies in the middle of the kinase *Pto* cluster (Salmeron et al., 1996). As, effectors can target different kinases, some NBS-LRRs can monitor multiple kinases (Lewis et al., 2013). Unlike direct recognitions, the indirect strategy potentially provides a wider protection against pathogens (Cesari, 2017) and may be favored in long-lived organisms. Guard components will be potentially modified by effectors in different ways (e.g., phosphorylation, acetylation, cleavage, etc.) and consequently detected by NBS-LRRs proteins as illustrated by the RIN4 guardee/RPS2-RPM1 CNLs model (Mackey et al., 2003). In this case, the only way for the pathogen to escape recognition is a loss of function of the effector, which may be insurmountable and lead to durable resistance.

## Expression and Regulatory Region

The predicted regulatory region of *Ma* is large and complex. Nevertheless, the transfer of *Ma* together with the 5.4 kb region localized immediately upstream the ATG in susceptible plants, was shown to be sufficient to confer the resistance (Claverie et al., 2011). The pattern of *cis*-acting binding sites is very similar between *RMjaGC2* and *Prupe.7G065400* putative promoter regions but quite different from the *Ma* region. The WRKY transcription factors, which bind to *cis*-regulatory W-boxes, respond to activation of abscisic acid-dependent pathway and belong to SA signaling that are both linked to NBS-LRR regulation (Eulgem et al., 2000). We found more W-boxes in the *Ma* regulatory region than in the *RMjaGC2* region. The *Prupe.7G065400* upstream region is clearly the least endowed with W-boxes. The insertion of the *Prupe.7G065500* gene that does not exist in plum and almond might disturb the regulatory region of the downstream *Prupe.7G065400*. Beyond the expression level, the sensitivity of the ON/OFF equilibrium of NBS-LRRs plays a crucial role to trigger or not the resistance. Indeed the switch toward an ON state of

the *L6/L7* alleles of the *L* locus is governed by the ligand affinity and the intramolecular constraints existing between TNL domains (Bernoux et al., 2016). These attributes defined how NBS-LRRs may cross the threshold conferring them the ability to produce, or not, cell death and/or resistance (Zhang et al., 2017).

We showed that *Ma* orthologs exhibit various and sizeable intronic microsatellites upstream each PL-coding exon with a particular feature in the last intron (PL4-PL5). This intriguing structural characteristic is present in 60% of the PL-containing TNLs in peach (Van Ghelder and Esmenjaud, 2016). We do not know yet if the presence of SSRs exclusively before PL-coding exons is only a consequence of genetic processes or if this structure plays a role in the modulation of TNL expression or splicing. Besides, several studies established a direct link between microsatellite size in intron and certain human diseases (Kersting et al., 2008; Suryadevara et al., 2013). Although the precise mechanisms are not fully understood, the microsatellite expansion seems to disrupt the targeted gene transcription (Bidichandani et al., 1998).

## The R Region Is a Combination of Dynamic and Preserved Loci

Transposable elements (TEs) are key factors in the birth-and-death model of R genes (Michelmore and Meyers, 1998). The increase of the KIN35-LRR25 interval size between peach and almond is due to TEs insertion. TEs disrupted the TNL environment inducing premature ends (*Prupe.7G065300*, *RMjaGC3*), quasi-total disruption (*RMjaGC1*) or NBS-LRR split (*Prupe.7G065500* - *Prupe.7G065600*) as also reported for *RMia* candidate genes in peach (Duval et al., 2014). The lectin tyrosine kinase is a highly preserved milestone contrasting with NBS-LRR cluster dynamism. Interestingly the kinase ortholog in Arabidopsis (*At2G41890*) lies in a region of the chromosome 2 that lacks the TNL cluster as the co-localization of the kinase and the *Ma* loci emerged in the *Prunoideae*. Our cluster study revealed two conserved TNL loci and other TNLs originated, disrupted or disappeared through recombination events. Evolution studies of R genes in the Arabidopsis genome proposed segmental and ectopic duplication as mode of NBS-LRR dispersion (Baumgarten et al., 2003; Leister, 2004). A ratio estimation of each mechanisms in three legume genomes revealed that local tandem duplication is the main mode of expansion (>75%) before ectopic or segmental duplications (Shao et al., 2014). Regardless of the type of mechanism, NBS-LRR duplications induce phenotypic variations through the diversification of sequences, and therefore in potential functionality, but also in gene expression (Wang et al., 2012). As illustrated in the *Ma* cluster, the modular structure of TNL genes together with a combination of partial duplications and deletions shape the TNL diversity. The chimeric *MaGC2* gene originated from an incomplete gene duplication of the *Ma* and *MaGC3* genes. Interestingly this gene includes a frameshift mutation close to the recombination site, making it prematurely truncated and perhaps preventing deleterious effects due to

auto activity. In this cluster, PL-coding exons are prone to intragenic duplications creating high sequence diversity. While *Prupe.7G065500* underwent PL2 duplication, the *Ma* orthologs underwent PL1 duplication followed by rapid PL3-PL4-PL5 duplications.

Finally, our mapping pointed out a restricted region conferring resistance to *M. javanica* in almond. The functional annotation, the genetic resources related to *Ma* and the peach genome, together with the transcript expression analysis strongly support that the *RMja* gene is the *Ma* ortholog. In a dynamic environment, the high conservation of the three orthologs, either conferring a complete-spectrum resistance (*Ma/plum*), or conferring an incomplete-spectrum resistance (*RMjaGC2/almond*), or conferring no resistance (*Prupe.7G065400/peach*), argues for a major role of these genes. Our study also identified the highest gene polymorphism in the atypical PL region that might be involved in the RKN resistance spectrum. These results pave the way to reveal the role of the PL domain and a putative original immune process for RKN control in *Prunus*.

## DATA AVAILABILITY STATEMENT

All relevant data is contain in the manuscript or in **supplementary material**, except the almond genome sequencing reads and the BAC clone sequences.

These datasets are publicly available in the NCBI BioProject accession number [PRJNA448729] [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA448729>] (BioSamples numbers [SAMN08863554] [<https://www.ncbi.nlm.nih.gov/biosample/SAMN08863554/>] and [SAMN08863555] [<https://www.ncbi.nlm.nih.gov/biosample/SAMN08863555/>]) and in the NCBI BioProject accession number [PRJNA448736] [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA448736>], respectively.

## AUTHOR CONTRIBUTIONS

HD, DE, and CVG designed the research project. HD, DE, MM, and CVG carried out and supervised the greenhouse and laboratory works. ED supervised and carried out the sequencing of almond genomes and produced the alignment files. CC supervised and carried out the almond genomic library, selected BAC clones and produced BAC clone sequences. HD supervised the data releases, CVG planned and designed the sequence analyses, analyzed the data, and drafted the manuscript. CVG and DE revised the manuscript. All authors commented and approved the manuscript.

## FUNDING

This work was supported by the French Ministry of Agriculture and Food [CASDAR project 'Pyrédune' n°C2011-01

(2011–2014)], the INRA BAP department (projects IVD2 and IVD3), the Plant-Nematode Interaction team (UMR ISA, Sophia Antipolis, France), and the GAFL Unit (INRA Montfavet, France).

## ACKNOWLEDGMENTS

The authors want to thank Emmanuel Botton, Laure Heurtevin, Nadia Therme, Ariane Fazari, and Ulysse Portier for their

contribution in tomato regeneration, molecular biology support, and phenotyping of almond progenies.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01269/full#supplementary-material>

## REFERENCES

- Baumgarten, A., Cannon, S., Spangler, R., and May, G. (2003). Genome-level evolution of resistance genes in *Arabidopsis thaliana*. *Genetics* 165, 309–319.
- Bernoux, M., Burdett, H., Williams, S. J., Zhang, X., Chen, C., Newell, K., et al. (2016). Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell* 28, 146–159. doi: 10.1105/tpc.15.00303
- Bidichandani, S. I., Ashizawa, T., and Patel, P. I. (1998). The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am. J. Hum. Genet.* 62, 111–121. doi: 10.1086/301680
- Blok, V. C., Jones, J. T., Phillips, M. S., and Trudgill, D. L. (2008). Parasitism genes and host range disparities in biotrophic nematodes: the conundrum of polyphagy versus specialisation. *Bioessays* 30, 249–259. doi: 10.1002/bies.20717
- Burch-Smith, T. M., Schiff, M., Caplan, J. L., Tsao, J., Czymbek, K., and Dinesh-Kumar, S. P. (2007). A novel role for the TIR domain in association with pathogen-derived elicitors. *PLoS Biol.* 5:e68. doi: 10.1371/journal.pbio.0050068
- Cesari, S. (2017). Multiple strategies for pathogen perception by plant immune receptors. *New phytol.* doi: 10.1111/nph.14877 [Epub ahead of print].
- Chalhoub, B., Belcram, H., and 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. doi: 10.1111/j.1467-7652.2004.00065.x
- Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474
- Chin, S. W., Shaw, J., Haberle, R., Wen, J., and Potter, D. (2014). Diversification of almonds, peaches, plums and cherries – Molecular systematics and biogeographic history of *Prunus* (*Rosaceae*). *Mol. Phylogenet. Evol.* 76, 34–48. doi: 10.1016/j.ympev.2014.02.024
- Claverie, M., Dirlwanger, E., Bosselut, N., Van Ghelder, C., Voisin, R., Kleinhentz, M., et al. (2011). The *Ma* gene for complete-spectrum resistance to *Meloidogyne* species in *Prunus* is a TNL with a huge repeated C-terminal post-LRR region. *Plant Physiol.* 156, 779–792. doi: 10.1104/pp.111.176230
- Duval, H., Hoerter, M., Polidori, J., Confolent, C., Masse, M., Moretti, A., et al. (2014). High-resolution mapping of the *RMia* gene for resistance to root-knot nematodes in peach. *Tree Genet. Genomes* 10, 297–306. doi: 10.1007/s11295-013-0683-z
- Eggermont, L., Verstraeten, B., and Van Damme, E. J. M. (2017). Genome-wide screening for lectin motifs in *Arabidopsis thaliana*. *Plant Genome* 10. doi: 10.3835/plantgenome2017.02.0010
- Elling, A. A. (2013). Major emerging problems with minor *Meloidogyne* species. *Phytopathology* 103, 1092–1102. doi: 10.1094/PHYTO-01-13-0019-RVW
- Esmenjaud, D., Voisin, R., Van Ghelder, C., Bosselut, N., Lafargue, B., Di Vito, M., et al. (2009). Genetic dissection of resistance to root-knot nematodes *Meloidogyne* spp. in plum, peach, almond, and apricot from various segregating interspecific *Prunus* progenies. *Tree Genet. Genomes* 5, 279–289. doi: 10.1007/s11295-008-0173-x
- Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206. doi: 10.1016/S1360-1385(00)01600-9
- Frith, M. C., Hansen, U., and Weng, Z. (2001). Detection of cis-element clusters in higher eukaryotic DNA. *Bioinformatics* 17, 878–889. doi: 10.1093/bioinformatics/17.10.878
- Gassmann, W., Hinsch, M. E., and Staskawicz, B. J. (1999). The *Arabidopsis* *RPS4* bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant J.* 20, 265–277. doi: 10.1046/j.1365-313X.1999.00600.x
- Gonthier, L., Bellec, A., Blassiau, C., Prat, E., Helmstetter, N., Rambaud, C., et al. (2010). Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., *Asteraceae*). *BMC Res. Notes* 3:225. doi: 10.1186/1756-0500-3-225
- Heath, M. C. (2000). Hypersensitive response-related death. *Plant Mol. Biol.* 44, 321–334. doi: 10.1023/A:1026592509060
- Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* 27, 297–300. doi: 10.1093/nar/27.1.297
- Holterman, M., Karssen, G., van den Elsen, S., van Megen, H., Bakker, J., and Helder, J. (2009). Small subunit rDNA-based phylogeny of the *Tylenchida* sheds light on relationships among some high-impact plant-parasitic nematodes and the evolution of plant feeding. *Phytopathology* 99, 227–235. doi: 10.1094/PHYTO-99-3-0227
- Janati, A., Berge, J. B., Triantaphyllou, A. C., and Dalmasso, A. (1982). Nouvelles données sur l'utilisation des isoestérases pour l'identification des *Meloidogyne*. *Fundam. Appl. Nematol.* 5, 147–154.
- Jaouannet, M., Magliano, M., Arguel, M. J., Gourgues, M., Evangelisti, E., Abad, P., et al. (2013). The root-knot nematode calreticulin Mi-CRT is a key effector in plant defense suppression. *Mol. Plant Microbe Interact.* 26, 97–105. doi: 10.1094/MPMI-05-12-0130-R
- Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Jones, J. T., Haegeman, A., Danchin, E. G. J., Gaur, H. S., Helder, J., Jones, M. G. K., et al. (2013). Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14, 946–961. doi: 10.1111/mpp.12057
- Karasov, T. L., Kniskern, J. M., Gao, L., DeYoung, B. J., Ding, J., Dubiella, U., et al. (2014). The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* 512, 436–440. doi: 10.1038/nature13439
- Kerstin, C., Agelopoulos, K., Schmidt, H., Korsching, E., August, C., Gosheger, G., et al. (2008). Biological importance of a polymorphic CA sequence within intron 1 of the epidermal growth factor receptor gene (EGFR) in high grade central osteosarcomas. *Genes Chromosomes Cancer* 47, 657–664. doi: 10.1002/gcc.20571
- Khallouk, S., Voisin, R., Portier, U., Polidori, J., Van Ghelder, C., and Esmenjaud, D. (2013). Multiyear evaluation of the durability of the resistance conferred by *Ma* and *Rmia* genes to *Meloidogyne incognita* in *Prunus* under controlled conditions. *Phytopathology* 103, 833–840. doi: 10.1094/PHYTO-09-12-0228-R
- Kochba, J., and Spiegel-Roy, P. (1975). Inheritance of resistance to the root-knot nematode (*Meloidogyne javanica* Chitwood) in bitter almond progenies. *Euphytica* 24, 453–457. doi: 10.1007/BF00028213
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends Genet.* 20, 116–122. doi: 10.1016/j.tig.2004.01.007

- Lewis, J. D., Lee, A. H., Hassan, J. A., Wan, J., Hurley, B., Jhingree, J. R., et al. (2013). The *Arabidopsis* ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the *Pseudomonas syringae* type III effector HopZ1a. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18722–18727. doi: 10.1073/pnas.1315520110
- Luo, S., Zhang, Y., Hu, Q., Chen, J., Li, K., Lu, C., et al. (2012). Dynamic nucleotide-binding site and leucine-rich repeat-encoding genes in the grass family. *Plant Physiol.* 159, 197–210. doi: 10.1104/pp.111.192062
- Mackey, D., Belkadir, Y., Alonso, J. M., Ecker, J. R., and Dangl, J. L. (2003). *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* 112, 379–389. doi: 10.1016/S0092-8674(03)00040-0
- Meyers, B. C., Dickerman, A. W., Micheltore, R. W., Sivaramakrishnan, S., Sobral, B. W., and Young, N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* 20, 317–332. doi: 10.1046/j.1365-3113X.1999.00606.x
- Micheltore, R. W., and Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8, 1113–1130. doi: 10.1101/gr.8.11.1113
- Milligan, S. B., Bodeau, J., Yaghoobi, J., Kaloshian, I., Zabel, P., and Williamson, V. M. (1998). The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. *Plant Cell* 10, 1307–1319. doi: 10.2307/3870642
- Narusaka, M., Kubo, Y., Hatakeyama, K., Imamura, J., Ezura, H., Nanasato, Y., et al. (2013). Breaking restricted taxonomic functionality by dual resistance genes. *Plant Signal. Behav.* 8:e24244. doi: 10.4161/psb.24244
- Nicol, J. M., Turner, S. J., Coyne, D. L., den Nijs, L., Hockland, S., and Maafi, Z. T. (2011). “Current nematode threats to world agriculture,” in *Genomics and Molecular Genetics of Plant–Nematode Interactions*, eds J. T. Jones, G. Gheysen, and C. Fenoll (Heidelberg: Springer), 21–44. doi: 10.1007/978-94-007-0434-3\_2
- Ntoukakis, V., Saur, I. M., Conlan, B., and Rathjen, J. P. (2014). The changing of the guard: the Pto/Prf receptor complex of tomato and pathogen recognition. *Curr. Opin. Plant Biol.* 20, 69–74. doi: 10.1016/j.pbi.2014.04.002
- Paal, J., Henselewski, H., Muth, J., Meksem, K., Menéndez, C. M., Salamini, F., et al. (2004). Molecular cloning of the potato *Gro1-4* gene conferring resistance to pathotype Ro1 of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. *Plant J.* 38, 285–297. doi: 10.1111/j.1365-3113X.2004.02047.x
- Peterson, D. G., Tomkins, J. P., Frisch, D. A., Wing, R. A., and Paterson, A. H. (2000). Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. *J. Agric. Genomics* 5, 1–100.
- Rairdan, G. J., and Moffett, P. (2006). Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *Plant Cell* 18, 2082–2093. doi: 10.1105/tpc.106.042747
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754
- Salmeron, J. M., Oldroyd, G. E., Rommens, C. M., Scofield, S. R., Kim, H. S., Lavelle, D. T., et al. (1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* 86, 123–133. doi: 10.1016/S0092-8674(00)80083-5
- Saucet, S. B., Van Ghelder, C., Abad, P., Duval, H., and Esmenjaud, D. (2016). Resistance to root-knot nematodes *Meloidogyne* spp. in woody plants. *New Phytol.* 211, 41–56. doi: 10.1111/nph.13933
- Shao, Z. Q., Zhang, Y. M., Hang, Y. Y., Xue, J. Y., Zhou, G. C., Wu, P., et al. (2014). Long-term evolution of nucleotide-binding site-leucine-rich repeat genes: understanding gained from and beyond the legume family. *Plant Physiol.* 166, 217–234. doi: 10.1104/pp.114.243626
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Sohn, K. H., Segonzac, C., Rallapalli, G., Sarris, P. F., Woo, J. Y., Williams, S. J., et al. (2014). The nuclear immune receptor RPS4 is required for RRS1SLH1-dependent constitutive defense activation in *Arabidopsis thaliana*. *PLoS Genet.* 10:e1004655. doi: 10.1371/journal.pgen.1004655
- Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 1, 10.1–10.12. doi: 10.1186/gb-2006-7-s1-s10
- Suryadevara, N. C., Neela, V. S., Devalraju, K. P., Jain, S., SivaSai, K. S., Valluri, V. L., et al. (2013). Influence of Intron II microsatellite polymorphism in human toll-like receptor 2 gene in leprosy. *Hum. Immunol.* 74, 1034–1040. doi: 10.1016/j.humimm.2013.04.009
- Tai, T. H., Dahlbeck, D., Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C., et al. (1999). Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14153–14158. doi: 10.1073/pnas.96.24.14153
- Tong, Z., Qu, S., Zhang, J., Wang, F., Tao, J., Gao, Z., et al. (2012). A modified protocol for RNA extraction from different peach tissues suitable for gene isolation and real-time PCR analysis. *Mol. Biotechnol.* 50, 229–236. doi: 10.1007/s12033-011-9433-3
- Truong, N. M., Nguyen, C. N., Abad, P., Quentin, M., and Favery, B. (2015). Function of root-knot nematode effectors and their targets in plant parasitism. *Adv. Bot. Res.* 73, 293–324. doi: 10.1016/bs.abr.2014.12.010
- van der Hoorn, R. A., and Kamoun, S. (2008). From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* 20, 2009–2017. doi: 10.1105/tpc.108.060194
- van der Vossen, E. A., Gros, J., Sikkema, A., Muskens, M., Wouters, D., Wolters, P., et al. (2005). The *Rpi-blb2* gene from *Solanum bulbocastanum* is an *Mi-1* gene homolog conferring broad-spectrum late blight resistance in potato. *Plant J.* 44, 208–222. doi: 10.1111/j.1365-3113X.2005.02527.x
- Van Ghelder, C., and Esmenjaud, D. (2016). TNL genes in peach: insights into the post-LRR domain. *BMC Genomics* 17:317. doi: 10.1186/s12864-016-2635-0
- Van Ghelder, C., Lafargue, B., Dirlwanger, E., Ouassa, A., Voisin, R., Polidori, J., et al. (2010). Characterization of the *RMja* gene for resistance to root-knot nematodes in almond: spectrum, location, and interest for *Prunus* breeding. *Tree Genet. Genomes* 6, 503–511. doi: 10.1007/s11295-010-0268-z
- Verde, I., Abbott, A. G., Scalabrini, S., Jung, S., Shu, S., Marroni, F., et al. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet.* 45, 487–494. doi: 10.1038/ng.2586
- Wang, Y., Wang, X., and Paterson, A. H. (2012). Genome and gene duplications and gene expression divergence: a view from plants. *Ann. N. Y. Acad. Sci.* 1256, 1–14. doi: 10.1111/j.1749-6632.2011.06384.x
- Whitham, S., Dinesh-Kumar, S. P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the *Tobacco mosaic virus* resistance gene *N*: similarity to toll and the interleukin-1 receptor. *Cell* 78, 1101–1115. doi: 10.1016/0092-8674(94)90283-6
- Wu, C. H., Abd-El-Halim, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., et al. (2017). NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 114, 8113–8118. doi: 10.1073/pnas.1702041114
- Zhang, J., Li, W., Xiang, T., Liu, Z., Laluk, K., Ding, X., et al. (2010). Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe* 7, 290–301. doi: 10.1016/j.chom.2010.03.007
- Zhang, X., Dodds, P. N., and Bernoux, M. (2017). What do we know about NOD-like receptors in plant immunity? *Annu. Rev. Phytopathol.* 55, 205–229. doi: 10.1146/annurev-phyto-080516-035250

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Van Ghelder, Esmenjaud, Callot, Dubois, Mazier and Duval. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.