BMC Genomics

RESEARCH Open Access

Comparative analysis of PRR receptors and R genes reveals structural and functional divergence between wild, ancestral, and commercial citrus species

Fredson dos Santos Menezes[1], Raner José Santana Silva[1], Patrick Ollitrault[2,3] and Fabienne Micheli[1,2,3*]

Abstract

Background Citriculture holds significant socioeconomic importance but is adversely affected by diseases that reduce productivity. Plant disease resistance is typically relies on efficient defense mechanisms - PTI and ETI - involving PRR receptors and R genes responsible for pathogen recognition and the initiation of defense responses.

Results In silico analyses performed on 11 genomes of 10 *Citrus* species, as well one genome of each *Poncirus trifoliata* and *Fortunella hindsii*, enabled the identification and classification of PRR receptors and R genes, analysis of gene duplication and clusters, validation of findings using RNA-seq libraries, and assessment of inter- and intraspecific variability across 43 additional accessions.

Conclusions Our results show that PRR receptors and genes are organized in clusters and exhibit both inter- and intraspecific diversity, with wild citrus species presenting lower gene counts compared to ancestral and hybrid species. Wild species showed a predominance of membrane-associated PRRs (RLKs and RLPs), potentially linked to early pathogen recognition, whereas ancestral and hybrid species exhibited higher abundance of intracellular receptors (CNL and TNL). These patterns suggest distinct immune strategies among citrus species, although further functional studies are required to confirm their roles in PTI or ETI responses.

Keywords Citriculture, Immune system, PRR receptors, R genes, Resistance clusters

Introduction

Citrus fruits belong to the Rutaceae family and are classified within the genera *Citrus*, *Poncirus*, *Fortunella*, *Microcitrus*, *Eremocitrus*, and *Clymenia*, with *Citrus* being the most commercially important [1]. Ancestral species of *Citrus* include *Citrus maxima*, *C. micrantha*, *C. medica*, and *C. reticulata* [2]. Genomic and phylogeographic studies suggest the center of origin for citrus species lies in the southwestern foothills of the Himalayas - specifically, the eastern region of Assam, northern Myanmar, and western Yunnan [2]. Today, citrus is cultivated in over 140 countries, predominantly in tropical

*Correspondence:
Fabienne Micheli
fabienne.micheli@cirad.fr
[1]Departamento de Ciências Biológicas (DCB), Centro de Biotecnologia e Genética (CBG), Universidade Estadual de Santa Cruz (UESC), Rodovia Ilhéus-Itabuna, Km 16, Ilhéus-BA 45662-900, Brazil
[2]UMR AGAP Institut, CIRAD, Montpellier 34398, France
[3]UMR AGAP Institut, CIRAD, INRAE, Univ Montpellier, Institut Agro, Montpellier, France

and subtropical zones between 35°N and 35°S, with most production concentrated in the northern hemisphere [3].

Citriculture holds substantial socioeconomic value, especially for the production of oranges and juices, generating employment across the entire production chain [4–6]. However, the crop faces major threats from diseases such Huanglongbing (HLB, caused by *Candidatus* Liberibacter spp.) [7], citrus canker (*Xanthomonas* subsp. *citri*) [8], citrus variegated chlorosis (*Xylella fastidiosa*) [9], citrus tristeza (*Closterovirus tristezae*) [10], citrus leprosis (*Cilevirus leprosis*) [11], citrus gummosis (*Phytophthora* spp.) [12] and brown spot of Alternaria (*Alternaria alternata f. sp. citri*) [13], which significantly reduce orchard productivity and growers' profits.

Plants possess a multilayered immune system that generally ensures resistance to pathogens[14, 15]. This system includes Pattern-Triggered Immunity (PTI), mediated by PRR receptors that detect pathogen-associated molecular patterns (PAMPs), and Effector-Triggered Immunity (ETI), which relies on intracellular R proteins that recognize pathogen effectors [15–17]. While these mechanisms share signaling components, they trigger distinct defense responses depending on recognition speed and specificity [18].

The emergence of PRR receptors and R genes has been associated with the birth-and-death model of gene evolution, driven by gene duplication events and likely co-evolving alongside phytopathogens [19, 20]. These genes undergo constant evolutionary pressures that result in polymorphisms related to copy number variation, intergenic exchanges, and allelic diversity [21]. Such variation arise mainly from duplication events, unequal crossing-over, sequence mutations, and selective pressure from pathogens interactions [22]. PRR receptors and R genes are broadly distributed across plant genomes and often found in clusters, typically formed by tandem duplications of paralogous sequences generated through unequal crossing-over. These duplications can expand gene families, and the resulting copies tend to diverge through mutation accumulation, increasing the complexity and specificity of plant immune responses [19, 23–26].

While the organization and clustering of these genes have been characterized in some species, including *C. sinensis* and *C. clementina* [27], broader studies are needed. A more comprehensive analysis should consider a wider range of gene classes, species, and varieties, as well as the relationship between gene distribution and the plant's defense capacity. Therefore, the objective of this work was to identify and characterize in silico PRR receptors and R genes in the genomes of *Citrus* spp., *Poncirus trifoliata*, and *Fortunella hindsii*, focusing on their structural and functional features, inter- and intraspecific variability, and their potential association with defense responses.

Material and methods

Input genomic data

The input data consisted of protein sequences in FASTA format from 11 genomes of *Citrus* species, one *Fortunella* genome, and one *Poncirus* genome, obtained from the Cluster Muse server (Cirad database, Meso@LR) and the Citrus Genome Database (https://www.citrusgenomedb.org/). The analyzed genomes included *Citrus reticulata*, *Citrus maxima*, *Citrus medica*, *Citrus micrantha* (data provided by Patrick Ollitrault – Cirad), *Citrus sinensis* cv. Valencia [28], *Citrus clementina* cv. Clemenules [29], *Citrus limon* Burm Primary [30], *Citrus limon* Burm Alternative [30], *Citrus ichangensis* (*C. cavaleriei*) cv. XJC [31], *Citrus australis* [32], *Poncirus trifoliata* [33], *Fortunella hindsii* S3y-45 v2.0 [34], *Atalantia buxifolia* cv. HKC [31] (Supplementary material 1).

Chromosome order correspondence

For *C. maxima*, *C. limon* Burm Primary, *C. limon* Burm Alternative, *C. sinensis*, *C. australis*, and *F. hindsii*, which present chromosomal assemblies in a different order compared to the reference genome of *C. clementina*, it was necessary to establish chromosome correspondence. Synteny relationships were determined using the Synteny View tool available at the Citrus Genome Database (https://www.citrusgenomedb.org/synview/search) (Supplementary material 2).

Domain annotation

Protein sequences from the selected genomes were subjected to domain annotation using InterProScan version 5.48–83.0 [35], executed on the Cluster Muse server (Meso@LR). The analysis was performed using default parameters, incorporating all available databases (COILS, Gene3D, HAMAP, MOBIDB, PANTHER, Pfam, PIRSF, PRINTS, ProDom, PROSITE, SFLD, SMART, SUPERFAMILY, and TIGRFAM). Output files were generated in Tab-Separated Values (TSV) format.

Identification and classification of R and PRR proteins

To identify and classify PRR and R proteins in the citrus genomes, the RRGPredictor tool was employed [36]. RRGPredictor has demonstrated higher efficiency compared to other available tools, enabling the detection of a broad range of PRR receptors and R genes, including those with variation in domain size and organization. The tool classifies proteins into 13 classes: T, TN, TNL, RPW8NL, RLP, RLK, RLKGNK2, N, NL, CN, CNL, MLO, and UNKNOWN. The classification was defined as follows:

- CN: Contains coiled-coil and nucleotide-binding site (NBS) domains, but lacks LRR.

- CNL: Contains coiled-coil, NBS, and leucine-rich repeat (LRR) domains.
- MLO: Contains an MLO-specific domain associated with mildew resistance.
- N: Contains only the NBS domain.
- NL: Contains NBS and LRR domains, but lacks N-terminal motifs.
- RLK: Contains an extracellular domain (e.g., LRR), a transmembrane domain, and an intracellular serine/threonine kinase domain.
- RLK-GNK2: Similar to RLK but with an additional GNK2 domain.
- RLP: Contains extracellular LRR domains and a transmembrane domain, but lacks a kinase domain.
- RPW8NL: Contains RPW8, NBS, and LRR domains.
- T: Contains only a Toll/interleukin-1 receptor (TIR) domain.
- TN: Contains both TIR and NBS domains.
- TNL: Contains TIR, NBS, and LRR domains.
- UNKNOWN: Contains immune-related domains but does not fit clearly into other classes.

Its high sensitivity and low operational cost make it suitable for identifying novel PRR and R gene candidates [36]. Chromosomal visualization of the identified genes was performed using TBtools tool [37].

In addition, the subcellular localization of all identified PRR and R proteins was predicted using the DeepLoc tool (version 2.1) (https://services.healthtech.dtu.dk/services/DeepLoc-2.1/). This tool classifies proteins based on their amino acid sequences into major cellular compartments such as the nucleus, cytoplasm, plasma membrane, endoplasmic reticulum, mitochondrion, and others. The predictions are provided in the Supplementary Material (Supplementary_material_15 to Supplementary_material_26), organized by gene class.

Gene duplication analysis

To assess gene duplication events, each proteome was first aligned against itself using the BLAST (Blastall) with the parameters: -p blastp -e 1e-10 -b 5 -v 5 -m 8. The resulting alignments were used to generate collinearity files for each species using MCScanX with default parameters. Duplication events or each class of identified PRR receptors and R genes were subsequently detected using the 'detect collinearity with gene families', also with default parameters [38]. This analysis produced duplication list and a graphical output for each gene class in each genome.

Identification of gene clusters

The identification of resistance gene clusters was performed using genomes with chromosome-level assemblies (i.e., *C. sinensis* cv. Valencia, *C. clementina*, *C. limon*, *C. australis, P. trifoliata*, and *F. hindsii*) and compared with previously obtained data from the four ancestral species (*C. maxima*, *C. micrantha*, *C. medica*, and *C. reticulata*; Submitted manuscript). Gene positions (i.e., start and end coordinates) were extracted from each species' GFF3 annotation files. Initially, IDs of all PRR receptors and R genes were extracted using the 'grep' command in the Linux terminal, with the genomic coordinates and chromosome locations. In the second step, these data were organized into Excel spreadsheets for each genome, creating individual sheets per chromosome and sorting genes by ascending genomic coordinates.

To identify gene clusters, we applied two complementary and widely used methodologies described by Holub and Jupe [39, 40]. First, Holub's method was used to identify clusters, defined as groups of three or more PRR or R located within 200 kb genomic window. This approach facilitates rapid detection, as it does not require the examination of intervening non-resistance genes. However, it does not capture clusters composed of only two genes. To address this limitation, Jupe's method was subsequently applied, which defines clusters as a group of two or more PRR or R genes within 200 kb, allowing for up to seven intervening non-NB-LRR genes. The combination of both methods enabled a comprehensive and robust identification of gene clusters across all examined genomes.

Validation using RNA-seq libraries

The expression analysis of PRR receptor and R gene transcripts was conducted using RNA-seq libraries retrieved from publicly available experiments deposited in the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra) for each species analyzed. Libraries derived from non-infected plants included: SRR22742835 (*C. australis*), SRR25461105 (*C. reticulata*), SRR15495635 (*C. medica*), SRR7614644 (*C. clementina*), SRR24101860 (*C. sinensis*), SRR18218357 (*C. limon*), SRR10356228 (*C. ichangensis*), SRR7614646 (*P. trifoliata*), and SRR9016287 (*F. hindsii*) (Supplementary material 3). Libraries derived from infected plants included: SRR23719598 (*C. maxima*), SRR14004272 (*C. medica*), SRR14728280 (*C. reticulata*), SRR14004267 (*C. sinensis*), SRR18218361 (*C. limon*), SRR10356230 (*C. ichangensis*), and SRR13480641 (*A. buxifolia*) (Supplementary material 3).

FASTQ files were downloaded using the Faster Download and Extract Reads in FASTQ tool on the Galaxy platform. Transcript expression (TPM) values were quantified using Salmon-Quant tool, referencing the transcriptome assemblies for each respective species: *C. reticulata*, *C. maxima*, *C. medica* (data from Cirad), *C. sinensis* cv. Valencia [28], *C. clementina* cv. Clemenules [29], *C. limon* Primary [30], *C. ichangensis* (*C. cavaleriei*) cv. XJC [31], *C. australis* [32], *Poncirus trifoliata* [33],

Fortunella hindsii S3y-45 v2.0 [34], *Atalantia buxifolia* cv. HKC [31].

Post-quantification, transcript data were filtered using the 'grep' command in Unix to retain only PRR and R gene entries. Final expression values were curated in Excel, and only genes with TPM > 0.0 and coverage ≥ 0.7 were considered as expressed.

Intraspecific diversity validation

Intraspecific diversity was assessed by comparing each of the four ancestral species with their respective accessions. Data specifically related to PRR receptors and R genes were extracted by filtering the results of a presence-absence variation (PAV) analysis conducted using the SGSGeneLoss tool [41], with the parameters cut-off = 0.9 and MinCoV = 5 (Submitted manuscript). This analysis compared the four ancestral species to 55 citrus accessions.

After filtering, the following accessions were selected: 13 for *C. maxima*, 16 for *C. reticulata*, 3 for *C. micrantha*, and 11 for *C. medica*. The presence and absence of PRR receptors and R genes were compared as follows:

- *C. maxima* was compared to Chandler pummelo (pum_cha2D), Haploid Chandler pummelo (pum_chaH), pummelo from Chongqing No016 (pum_ChonD), Flores pummelo (pum_florD), pummelo from Guilin No1 (pum_GuilD), F22 pummelo (pum_hf22D), Huanonghongyou (pum_HuanD), Huazhoujuhong (Pum_HuazD), Indian pummelo (pum_indeD), Kao pan pummelo (pum_kaopD), Reinking pummelo (pum_reinD), Tahiti pummelo (pum_tahiD), and Timor pummelo (pum_timoD).
- *C. reticulata* was compared to Avana Tardivo di Ciaculli mandarin (man_avanD), Cleopatra mandarin (man_clo1D), Daoxian wild mandarin No1 (man_DaoW1D), *Citrus daoxianensis* (man_daoxD), Hezhou wild mandarin (man_HezWD), Jiangyong wild mandarin (man_JiaWD), King of Siam mandarin (man_kingD), MS2 mandarin (man_ManWD), Nan feng mi chu mandarin (man_nfmcD), Owari mandarin (man_owarD), Ponkan mandarin (man_ponkD), Sun chu sha mandarin (man_scshD), Shekwasha mandarin (man_shekD), Sunki mandarin (man_sunkD), Szinkom mandarin (man_szinD), and Tachibana mandarin (man_tachD).
- *C. micrantha* was compared to *C. macroptera* (pap_macrD), Melanesian papeda (pap_melaD), and Biasong papeda (pap_mic1D).
- *C. medica* was compared to Buddhas hand citron (cit_BdHD), Chinese citron (cit_chinD), Corsican citron (cit_cor1D), Diamante citron (cit_diamD), Etrog-861 citron (cit_e861D), Humpang citron (cit_humpD), Mac Nao San citron (cit_macnD), Mac Veu Mountain citron (cit_macvD), wild citron1 (cit_wc1D), wild citron2 (cit_wc2D), and wild citron3 (cit_wc3D).

Phylogenetic inference of RLK genes associated with resistance traits

To explore the evolutionary relationships and potential roles in RLK genes in citrus immunity, a phylogenetic analysis was performed using RLK proteins sequences identified in 11 citrus genomes: *C. reticulata, C. maxima, C. medica, C. micrantha, C. sinensis, C. clementina, C. limon, C. ichangensis, A. buxifolia, P. trifoliata, and F. hindsii.* The RLK sequences were aligned using MAFFT v7.475 [42] with default parameters, and poorly aligned regions were trimmed using TrimAl [43]. Phylogenetic trees were inferred using PhyML [44], applying the following parameters: *phyml -i RLK_aligned.phy -d aa -m LG -c 4 -a e*, using Bayesian inference and maximum likelihood. Phylogenetic trees were visualized and edited using the InTreeGreat tool (https://phylogeny.southgreen.fr/treedisplay/index.php). Orthologous RLK gene groups were identified using the Tree Pattern Matching tool available in the Citrus Phylogeny Database, developed in collaboration with Dr. Jean-François Dufayard (Cirad, France) associated with resistance traits in citrus species.

Results

Interspecific diversity of PRR receptors and R genes

Classification and quantitative comparison among citrus species

PRR receptors and R genes were categorized into 13 classes: CN, CNL, ML, N, NL, RLK, RLKGNK2, RLP, RPW8NL, T, TN, TNL, and UNKNOWN. Their abundance was quantified and compared across eight citrus genomes and four ancestral species (Fig. 1; Supplementary material 4).

Ancestral species (Fig. 1, red rectangle) exhibited the highest total numbers of PRR and R genes. In comparison, the three commercially important hybrid species - *C. sinensis*, *C. clementina*, and *C. limon* (Fig. 1, green rectangle) - showed a moderate reduction in receptor abundance relative to their ancestral lineages. Conversely, these hybrids presented higher receptor counts than the wild species (Fig. 1, blue rectangle), which displayed the lowest totals overall.

Species from the genera *Fortunella* and *Poncirus* (Fig. 1, orange rectangle) showed intermediate profiles. *Fortunella* species were more similar to ancestral genomes, whereas *Poncirus* aligned more closely with hybrid species in terms of gene abundance.

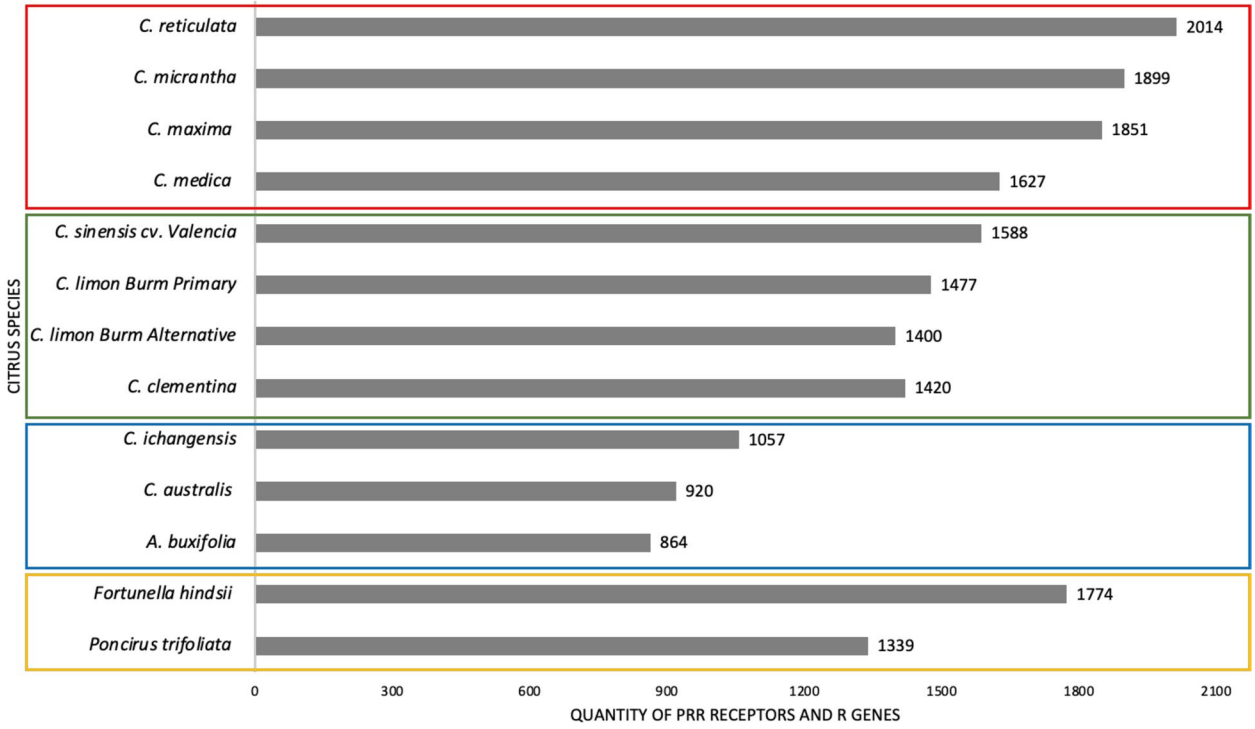

Fig. 1 Comparison of the total number of PRR receptors and R genes present in ancestral species (highlighted in the red rectangle) with the descendant "hybrid" species, *C. sinensis* cv. Valencia, *C. limon* Burm Primary, *C. limon* Burm Alternative, and *C. clementina* (highlighted in the green rectangle); wild species *C. ichangensis*, *C. australis*, and *A. buxifolia* (highlighted in the blue rectangle); and species belonging to other genera, *F. hindsii* and *P. trifoliata* (highlighted in the yellow rectangle)

Diversity patterns of PRR and R gene classes among citrus genomes

The relative distribution of PRR receptor and R genes (excluding the UNKNOWN class) across citrus genomes revealed that CNL, NL, RLK, and TNL were the predominant classes in most species, with the exception of *C. ichangensis*, *C. australis*, and *A. buxifolia*. Notably, the TNL class was not among the most significant in *C. maxima* (Fig. 2). In *C. australis*, membrane-localized receptor classes - RLK, RLKGNK2, and RLP - were the most represented. Similarly, RLKGNK2 and RLP were dominant in *C. ichangensis*, while *A. buxifolia* showed a marked prevalence of RLK genes (Fig. 2).

These wild citrus species exhibited the lowest total number of genes encoding the Tir domain (i.e., T, TN, and TNL classes). Specifically, *C. ichangensis* contained only six TIR-domain genes (four T, one TN, and one TNL), *C. australis* had 13 (12 T and one TN, with no TNL), and *A. buxifolia* had 24 (11 T, three TN, and 10 TNL) (Supplementary material 4).

Chromosomal distribution and gene duplication patterns

General chromosomal distribution across citrus genomes

In *C. limon* and the ancestral *C. maxima*, PRR and R genes were predominantly located on chromosomes 2, 5, and 7 (Fig. 3). Although *C. australis* exhibited a lower total number of identified receptors, it showed a more even chromosomal distribution, with higher gene counts on chromosomes 1, 2, 3, and 7 (Fig. 3). To visualize the chromosome organization of the 12 PRR and R gene classes and to highlight potential duplication events, circos-style genome representations were generated for species with chromosome-level assemblies (Fig. 4).

Class-specific chromosomal distribution of CNL, TNL, and RLK genes

In the CNL class, *C. clementina*, *C. limon, P. trifoliata*, and *F. hindsii* exhibited distributions patterns similar to the ancestral species *C. reticulata* and *C. medica*, with higher gene counts on chromosome 5. In contrast, *C. sinensis* showed a pattern more similar to *C. micrantha* and *C. maxima*, with a greater concentration on chromosome 7. For the TNL class, all descendant citrus species, as well *P. trifoliata*, *F. hindsii*, and all ancestral species, presented higher gene densities on chromosome 3. Regarding the RLK class, *C. clementina*, *C. sinensis*, *P. trifoliata*, and *F. hindsii* resembled *C. medica* and *C. micrantha*, with most genes located on chromosome 1. Conversely, *C. limon* mirrored the distributions observed in *C. reticulata* and *C. maxima*, with higher gene counts on chromosome 2. Notably, the wild species *C. australis* exhibited a distinct pattern characterized by the

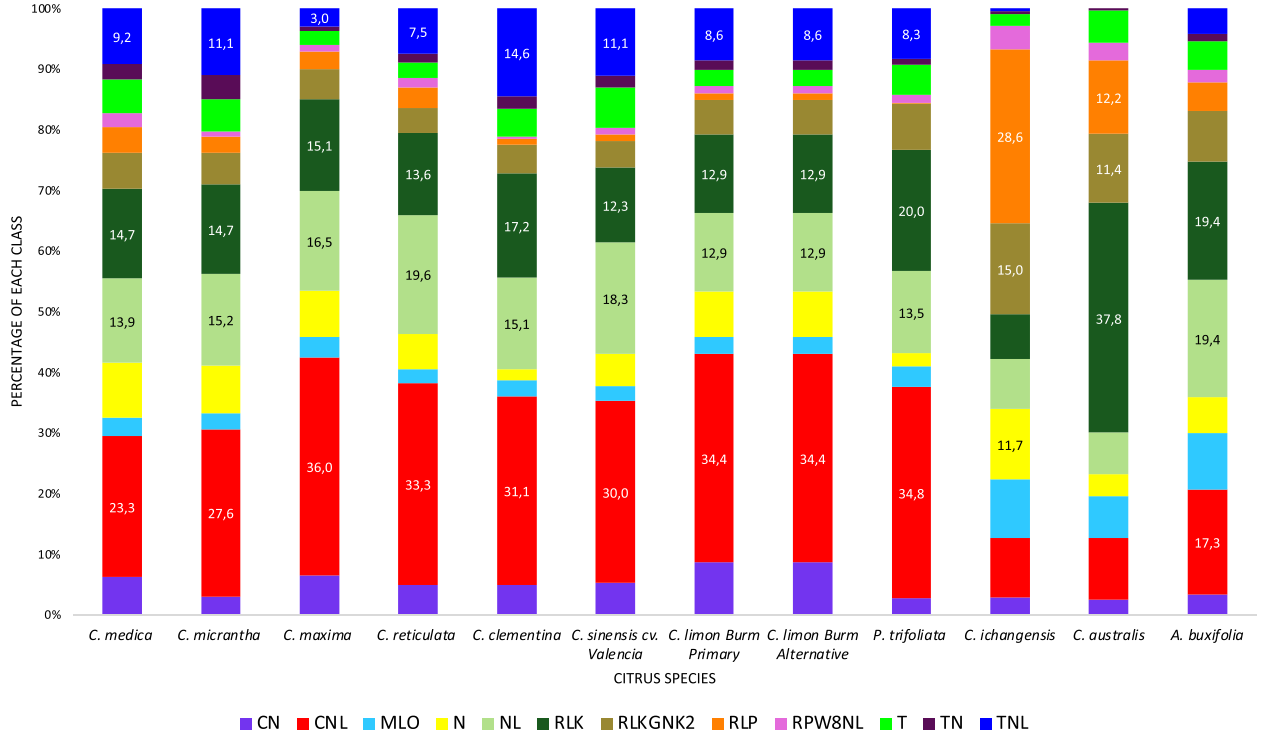

Fig. 2 Comparative analysis of the distribution of CN, CNL, ML, N, NL, RLK, RLKGNK2, RLP, RPW8NL, T, TN, and TNL classes in the genomes of *C. medica*, *C. micrantha*, *C. maxima*, *C. reticulata*, *C. clementina*, *C. sinensis*, *C. limon* Burm Primary, *C. limon* Burm Alternative, *P. trifoliata*, *F. hindsii*, *C. ichangensis*, *C. australis*, and *A. buxifolia species*

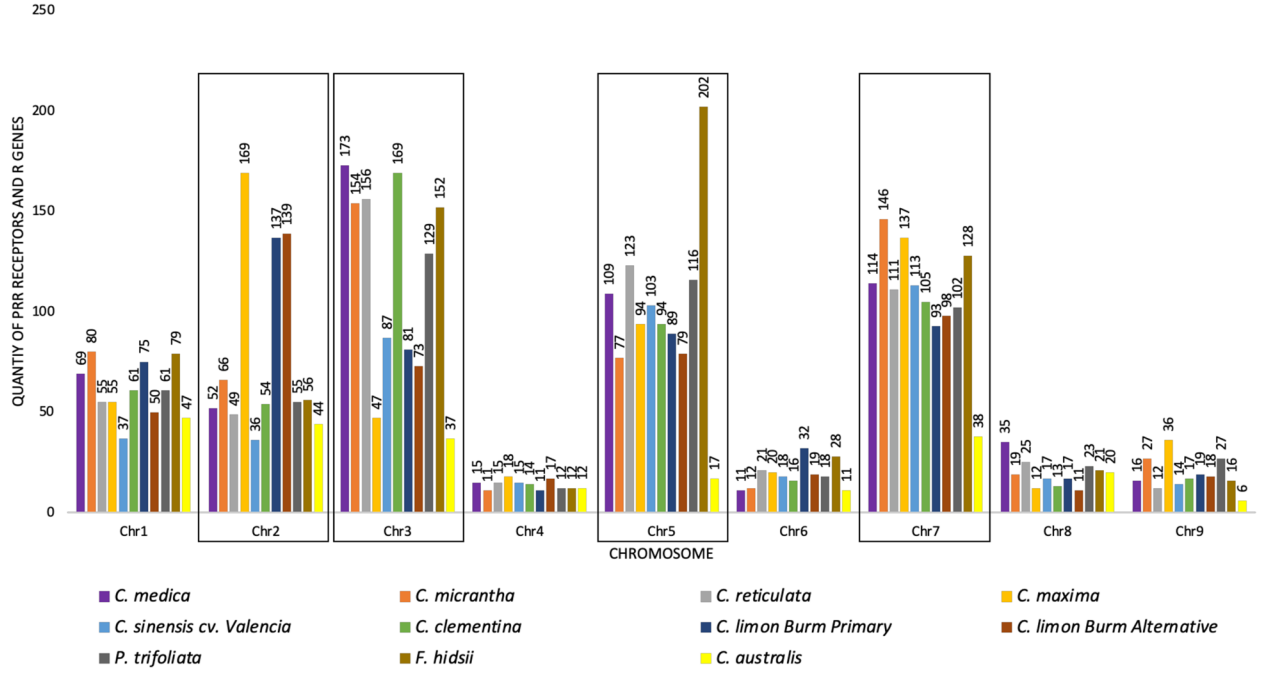

Fig. 3 Comparative analysis of the chromosomal distribution of the quantity of genes belonging to the CN, CNL, ML, N, NL, RLK, RLKGNK2, RLP, RPW8NL, T, TN, and TNL classes in the species *C. medica*, *C. micrantha*, *C. maxima*, *C. reticulata*, *C. clementina*, *C. sinensis*, *C. limon* Burm Primary, *C. limon* Burm Alternative, *P. trifoliata*, *F. hindsii*, and *C. australis*

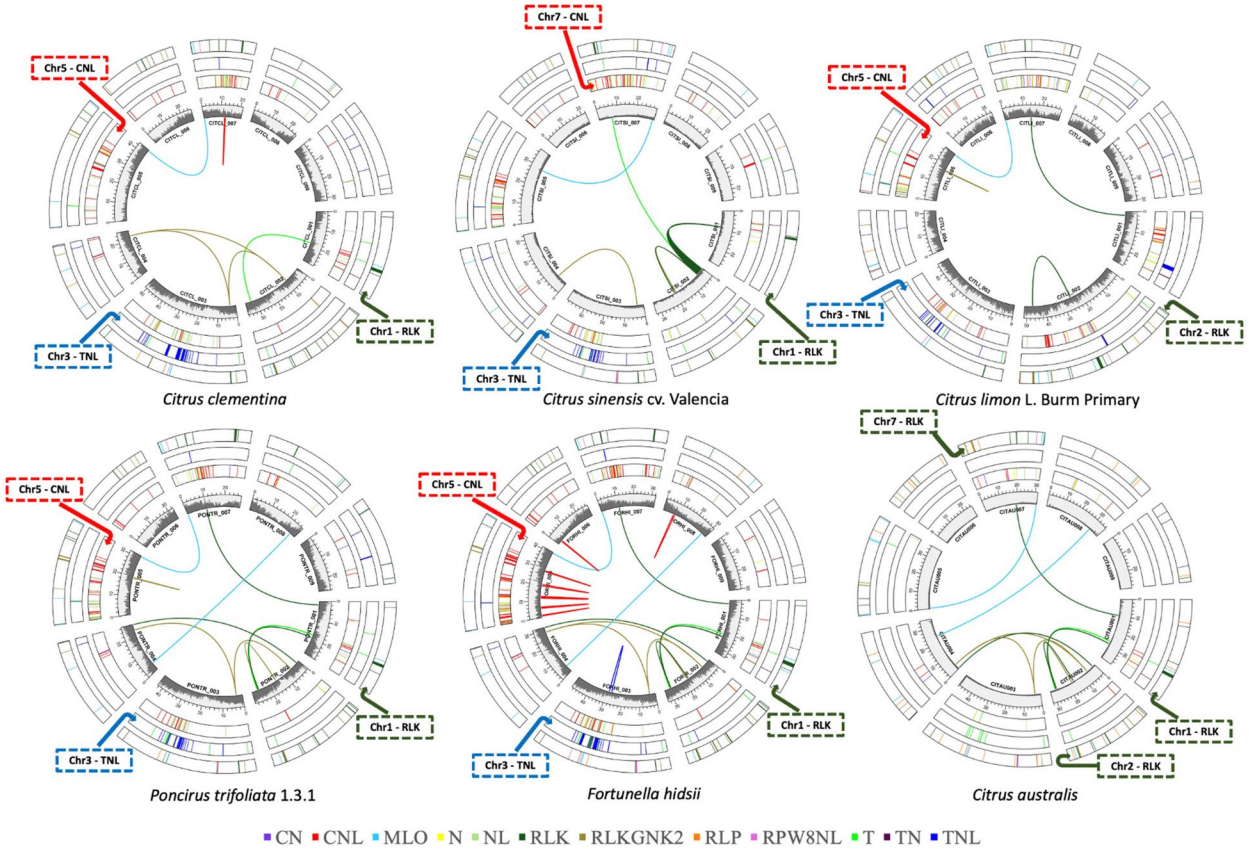

Fig. 4 Circos-style representation of the distribution of PRR receptors and R genes across the nine chromosomes of the species *C. clementina*, *C. sinensis*, *C. limon* Burm Primary, *P. trifoliata*, *F. hindsii*, and *C. australis*. Duplication events are represented by lines in the central inner part of the circos, with respective colors representing the class. Dashed red rectangles indicate the chromosome with the highest quantity of CNL receptors, dashed green rectangles indicate the chromosome with the highest quantity of RLK receptors, and dashed blue rectangles indicate the chromosome with the highest quantity of TNL receptors

predominance of the RLK class across all chromosomes, particularly chromosomes 1, 2, and 7 (Supplementary material 5).

Gene duplication dynamics in the RLK class

Gene duplication events within the RLK class were detected across all analyzed species (Supplementary material 6). Duplications on chromosomes 1 and 2 were observed in all species except *C. clementina*. Conversely, *C. maxima*, *C. clementina*, *C. sinensis*, and *C. limon* did not exhibit RLK duplications on chromosome 4 (Supplementary material 7). The wild species *C. australis*, *P. trifoliata*, and *F. hindsii* showed the greatest diversity in duplicated RLK genes, with events spanning chromosomes 1, 2, 4, and 7.

In six of the studied species, one or more genes from the RLK genes appeared to originate from common ancestral genes. In *C. medica*, for instance, CITME_002t027420 and CITME_002t018330 were duplicated from CITME_001t015860; CITME_002t27490 and CITME_001t015860, from CITME_002t018330; and CITME_001t015780 and CITME_022t018300, from CITME_002t027430. In *C. micrantha*, CITMI_001t015480 and CITMI_002t4730 were duplicated from CITMI_002t28800. In *C. maxima*, CITMA_002t014590 and CITMA_002t22630 originated from CITMA_002t041020. In *C. reticulata*, CITRE_001t014640 and CITRE_002t013780 were duplicated from CITRE_002t27370. In *C. sinensis*, Cs4g11660.1 and Cs4g11630.1 were duplicated from Cs2g7100, while Cs4g22390.1 and Cs4g11730.1 originated from Cs207160.1. In *C. australis*, g19680.t1 and g22278.t1 were duplicated from g23618.t1 (Supplementary material 7).

Genomic organization of PRR and R gene clusters

Clustering patterns and chromosomal hotspots

The ancestral citrus species exhibited a higher average percentage of clustered of PRR and R genes compared to the other species analyzed, with *C. australis* showing the lowest clustering rate (72.7%, Table 1). Chromosomal distribution of these clusters varied among species: chromosome 2 was the major hotspot in *C. australis* and *C. limon*, and the ancestral *C. maxima*; chromosome 3

Table 1 Comparative analysis of clustering structure in ancestral species *C. maxima*, *C. micrantha*, *C. medica*, and *C. reticulata* with *F. hindsii*, *C. limon* Burm Primary, *C. clementina*, *P. trifoliata*, *C. sinensis*, and *C. australis* species, representing the number of clusters (NC) and the percentage of genes in clusters on each chromosome (chr), as well as the total number of clusters and the overall percentage of genes in clusters. The color scale varies from green to red according to the highest to lowest percentages of genes in clusters. Chromosomes with the highest number of clusters in a given species are shown in bold

Espécies	chr1		chr2		chr3		chr4		chr5		chr6		chr7		chr8		chr9		Total	
	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*	NC	% de gene em *clusters*
C. maxima	32	84,7	**72**	90,6	31	70,0	20	81,4	46	90,2	23	81,9	58	94,3	23	93,0	27	82,0	332	87,2
C. micrantha	31	86,8	40	87,6	**58**	88,5	18	73,2	43	85,6	15	72,1	55	95,3	16	75,3	26	84,7	302	86,8
C. medica	35	87,2	33	79,0	59	89,8	21	74,1	52	94,0	16	70,8	**60**	92,6	19	78,6	19	74,7	314	85,9
C. reticulata	29	84,4	32	80,1	**59**	86,5	20	77,2	50	92,8	16	73,2	46	89,8	20	81,3	26	86,8	298	85,5
F. hindsii	30	83,8	37	80,6	55	82,7	20	74,7	**65**	92,8	15	65,5	53	90,2	19	81,4	21	80,6	315	84,3
C. limon	28	87,2	**59**	87,3	38	78,8	13	57,7	31	86,8	17	76,3	48	90,7	12	88,9	16	76,3	262	84,0
C. clementina	24	82,5	34	79,1	**59**	89,1	22	79,8	40	83,0	13	71,3	36	93,4	11	66,7	18	78,8	257	83,6
P. trifoliata	26	81,3	34	79,6	45	81,7	18	69,1	37	86,5	16	70,9	**47**	92,7	23	87,5	23	83,5	269	83,2
C. sinensis	17	77,0	32	82,6	37	78,1	21	68,5	**45**	90,2	19	81,2	**45**	88,7	13	75,5	16	76,8	245	81,8
C. australis	27	80,1	**31**	80,7	25	64,6	19	74,6	20	66,3	12	61,6	29	79,9	14	76,5	6	40,9	183	72,7

in *C. clementina*, and the ancestral *C. reticulata* and *C. micrantha*; chromosome 5 in *F. hindsii* and *C. sinensis*; and chromosome 7 in *P. trifoliata* and *C. sinensis*, and the ancestral *C. medica*. In *C. sinensis*, clusters were equally distributed between chromosomes 5 and 7 (Table 1).

Given their chromosomal distribution, presence of duplications in all species, and functional importance in PAMP and/or effector recognition, RLK gene clusters were selected for more detailed investigation on their predominant chromosomes (Supplementary material 8). The species with the highest number of RLK genes organized in clusters on a single chromosome were *C. maxima* (55 genes in 13 clusters), *F. hindsii* (44 genes in 8 clusters), and *C. limon* (39 genes in 8 clusters). In contrast, *C. reticulata* and *C. sinensis* presented the lowest counts, with 18 and 23 clustered RLK genes forming five and four clusters, respectively. These formed five clusters in *C. reticulata* and four in *C. sinensis*. The largest individual RLK clusters were found in *F. hindsii* and *P. trifoliata*, containing 15 and 14 genes, respectively (Supplementary material 8).

Selection and characterization of representative PRR and gene clusters

For detailed investigation, representative clusters were selected from predominant chromosomes based on two criteria: (i) those containing the highest number of genes in each species (designated as Cluster A), and (ii) those containing genes with two or more duplication events (designated as Clusters B) (Supplementary material 9). After filtering for clusters enriched in RLK genes, the following were selected: one Cluster A and one Cluster B in *C. maxima*, *C. reticulata*, *C. micrantha*, *C. sinensis*, and *C. australis*; one Cluster A and two Cluster B in *C. medica*; and one Cluster A only in *C. clementina*, *C. limon*, *P. trifoliata* and *F. hindsii*.

Within these clusters, genes with two or more duplication events are highlighted in bold green, while those genes involved in a single duplication event are marked in yellow (Supplementary material 9). Duplication events were identified in Cluster A of *C. maxima*, *C. medica*, *C. micrantha*, and *C. sinensis*. Notably, in *C. medica*, exhibited both single and multiple duplication events within the same cluster.

The presence of genes from other PRR or R gene classes within RLK-enriched clusters was also noted: *C. maxima* (one UNKNOWN class gene), *C. medica* (one N, one NL, and one UNKNOWN class gene); in *C. sinensis* (one UNKNOWN class gene); *C. limon* (two UNKNOWN and one N class gene); and *P. trifoliata* (two UNKNOWN and one CN class gene).

In Cluster B, genes with two or more duplication events were observed in *C. maxima*, *C. reticulata*, *C. micrantha*, and *C. australis*; two such genes were found in *C. sinensis*, and one in each of the two *C. medica* Cluster B regions (along with another gene involved in a single event). Additionally, clusters B in *C. maxima*, *C. medica*, and *C. reticulata* contained insertions of RLKGNK2 genes - two in the clusters of *C. medica* and *C. reticulata*, and three in *C. maxima* (Supplementary material 9).

Expression profiles of PRR and R genes

Baseline expression in healthy leaf tissues

RNA-seq libraries derived from citrus leaf tissues not exposed to pathogens were analyzed to assess the baseline expression of PRR and R genes (Table 1). The proportion of expressed transcripts among identified genes ranged from 64.3% in *C. ichangensis* to 79.4% in *F. hindsii*.

For each species, we identified the predominance PRR and R gene classes based on the average transcript abundance (TPM), within their respective libraries. For example, in *C. australis* transcripts from the CNL, N, NL, RLK, RLP, and RPW8NL classes were the most abundant. In other species, MLO, RLKGNK2, and RPW8NL classes were among those with the highest expression within their respective datasets (Table 2).

Quantitative expression profiles across RNA-seq libraries during pathogen infection

RNA-seq libraries from citrus leaves infected by phytopathogens were analyzed to evaluate the expression patterns of PRR and R genes under biotic stress conditions (Table 2). The proportion of expressed transcripts among identified genes ranged from 58.9% in *C. ichangensis* to 82.4% in *C. medica*.

For each species, the PRR and R gene classes with highest TPM within their respective libraries were identified. For example, in *C. medica*, transcripts from CN, CNL, RLK, and RPW8NL classes were the most abundant. In *C. reticulata*, the MLO, RLKGNK2, RLP, and RPW8NL classes showed higher expression within its dataset. *Additionally,* in *C. maxima*, the MLO and RLKGNK2 classes were prominent; in *C. sinensis,* the RPW8NL class was abundant; in *C. limon,* the MLO class predominated; and in *A. buxifolia* and *C. ichangensis*, the RLKGNK2 and RPW8NL classes were among the most expressed (Table 2).

Expression profiles of RLK-a and RLK-b clusters in healthy and infected libraries

For RNA-seq libraries derived from uninfected citrus leaves, all transcripts from RLK clusters A and B were expressed in *C. australis* and *C. reticulata*. In both species, cluster B, although containing fewer genes, exhibited higher average TPM values than cluster A (Table 3).

In RNA-seq libraries obtained from pathogen-infected tissues, all transcripts from RLK clusters A and B were expressed in *C. medica* and *C. sinensis*. In *C. maxima*, *C. reticulata*, and *C. sinensis*, cluster B, which includes genes involved in two or more duplication events, showed higher average TPM values than to cluster A within each species' dataset (Table 3).

Intraspecific variation of PRR and R genes

Gene copy number variation among citrus accessions

Gene copy variation was assessed for PRR and R genes across accessions within ancestral citrus species. Among *C. maxima* and 13 additional pomelo accessions, the standard deviation in total gene count of 46.01. The highest variability was observed in the CNL and NL classes, with standard deviations of 25.59 and 10.09, respectively. In contrast, the RLK class showed limited variation (standard deviation = 1.94), while the RPW8NL class displayed fully conserved gene counts across all accessions.

In *C. reticulata* and 16 other mandarin accessions, a higher overall standard deviation of 64.16 was observed. the CNL, NL, and RLK classes exhibited the greatest deviations (25.01, 15.75, and 6.72, respectively), whereas the MLO and RPW8NL classes remained highly conserved, with deviations of 0 and 0.24, respectively.

Among *C. micrantha* and three additional papeda accessions, the highest overall variation among the ancestral species was detected, with a standard deviation of 108.39. The CNL, NL, and TNL classes contributed most to this variability (30.73, 19.17, and 14.24, respectively), while the RLK class also showed moderate variation (standard deviation = 6.24). In contrast, the MLO and RPW8NL classes showed conserved gene counts across accessions.

In *C. medica* and 11 additional citron accessions, gene copy numbers were more conserved overall, with a standard deviation of 16.77. All gene classes exhibited low variability, with the highest standard deviations found in the TNL (3.44) and NL (3.33) classes. The MLO and RPW8NL classes were fully conserved or nearly so (standard deviations of 0 and 0.29, respectively).

Conservation of RLK genes clusters among citrus accessions

Analysis of RLK gene clusters across citrus accessions revealed differential patterns of conservation. In *C. maxima* and *C. micrantha*, the identified and selected clusters remained fully conserved in terms of gene presence across all analyzed accessions. In contrast, variation was observed in the RLK clusters of *C. reticulata* and *C. medica*.

In *C. reticulata*, the gene CITRE_002g001910 (cluster A) was absent in six accessions: MS2 mandarin, *C. daoxianensis*, Hezhou wild mandarin, Ponkan mandarin, Shekwasha mandarin, and Sunki mandarin. Additionally, the gene CITRE_002g002060 was not detected in the Avana Tardivo di Ciaculli accession, while both genes were absent in the Jiangyong wild mandarin accession. All other accessions retained complete gene content for cluster A.

In *C. medica*, the gene CITME_001g015660 (cluster A) was not detected five accessions: Mac Veu Mountain citron, wild citron2, Mac Nao San citron, Humpang citron,

Table 2 Validation of the expression of PRR receptors and R genes in RNAseq libraries of leaves not infected and infected by some type of pathogen

	Species	RNAseq libraries	Expression level	CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	UNKNOWN	Average
RNAseq libraries from leaves not infected by pathogens	*Citrus australis*	SRR22742835	% of expressed transcripts	83.3	84.0	70.6	77.8	100.0	98.9	85.7	43.3	28.6	15.4	100.0	0.0	69.3	71.9
			Average TPM of expressed PRRs and R genes	1.6	16.0	16.6	18.7	12.4	15.3	9.7	15.9	17.3	1.5	0.1	-	14.4	11.6
	Citrus reticulata	SRR25461105	% of expressed transcripts	51.4	89.2	88.2	38.6	85.7	87.3	80.0	15.4	16.7	47.4	72.7	96.4	61.5	67.9
			Average TPM of expressed PRRs and R genes	3.0	4.8	16.4	2.5	4.3	8.1	15.6	10.8	16.0	7.6	0.4	5.0	11.2	8.1
	Citrus medica	SRR15495635	% of expressed transcripts	64.9	92.8	83.3	42.6	80.7	94.3	86.1	24.0	7.1	39.4	93.3	94.5	68.1	71.2
			Average TPM of expressed PRRs and R genes	1.4	4.9	16.0	1.7	4.9	9.7	4.8	2.6	11.4	8.9	2.6	3.6	6.7	6.1
	Citrus clementina	SRR7614644	% of expressed transcripts	85.7	92.8	73.9	45.5	82.2	85.6	75.8	33.3	100.0	34.5	72.7	91.0	75.6	78.6
			Average TPM of expressed PRRs and R genes	5.8	5.7	6.5	3.6	4.0	3.0	8.0	4.4	11.5	4.8	0.5	2.2	9.3	5.3
	Citrus sinensis	SRR24101860	% of expressed transcripts	44.1	93.7	66.7	38.5	83.3	84.5	88.2	0.0	28.6	23.3	53.9	89.2	67.8	71.7
			Average TPM of expressed PRRs and R genes	4.2	5.6	8.3	5.4	4.7	7.7	13.7	-	8.9	3.0	1.2	4.4	10.5	6.5
	Citrus limon	SRR18218357	% of expressed transcripts	64.3	96.5	72.2	20.4	87.1	95.3	84.2	14.3	25.0	11.1	70.0	91.4	67.4	73.1
			Average TPM of expressed PRRs and R genes	2.9	2.8	16.2	3.0	1.7	4.3	2.2	0.0	3.1	7.6	3.0	3.7	5.5	4.3
	Citrus ichangensis	SRR10356228	% of expressed transcripts	66.7	91.3	59.1	33.3	83.3	83.3	75.7	25.4	22.2	40.0	100.0	100.0	55.9	64.3
			Average TPM of expressed PRRs and R genes	9.7	7.1	8.6	4.7	10.0	8.9	12.7	5.3	11.0	1.3	0.0	4.2	10.2	7.2
	Poncirus trifoliata	SRR7614646	% of expressed transcripts	80.0	87.3	63.6	69.2	89.5	83.6	56.5	0.0	14.3	29.6	80.0	62.7	74.9	75.7
			Average TPM of expressed PRRs and R genes	6.9	4.7	12.1	2.2	3.6	4.1	7.5	-	11.6	5.3	1.1	3.7	9.9	6.1
	Fortunella hindsii	SRR9016287	% of expressed transcripts	53.3	88.3	75.0	46.2	85.8	87.2	88.2	25.0	23.1	48.1	85.7	90.2	76.9	79.4
			Average TPM of expressed PRRs and R genes	2.4	7.1	24.4	6.2	5.2	8.2	14.8	2.4	12.1	3.8	5.8	17.8	14.1	9.6
	Species	**RNAseq libraries**		**CN**	**CNL**	**MLO**	**N**	**NL**	**RLK**	**RLKGNK2**	**RLP**	**RPW8NL**	**T**	**TN**	**TNL**	**UNKNOWN**	**Average**

Table 2 (continued)

	Species	RNAseq libraries	Expression level	CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	UNKNOWN	Average
RNAseq libraries of leaves infected by some type of pathogen	*Citrus maxima*	SRR23719598	% of expressed transcripts	80.5	96.0	66.7	39.6	82.7	89.5	68.8	22.2	28.6	50.0	100.0	100.0	64.6	70.3
			Average TPM of expressed PRRs and R genes	5.0	6.4	32.0	4.0	5.1	6.8	15.6	2.3	7.8	4.6	2.2	9.7	10.4	8.6
	Citrus medica	SRR14004272	% of expressed transcripts	89.2	100.0	83.3	48.2	97.6	98.9	91.7	36.0	14.3	48.5	93.3	100.0	80.6	82.4
			Average TPM of expressed PRRs and R genes	18.3	14.9	11.1	5.3	8.9	14.0	12.1	1.3	20.8	4.3	8.8	11.2	11.2	10.9
	Citrus reticulata	SRR14728280	% of expressed transcripts	59.5	94.0	82.4	43.2	87.8	87.3	80.0	19.2	16.7	52.6	72.7	92.9	64.1	70.4
			Average TPM of expressed PRRs and R genes	5.2	8.7	17.5	2.2	7.9	8.2	16.3	17.8	26.8	6.2	1.0	4.5	13.0	10.4
	Citrus sinensis	SRR14004267	% of expressed transcripts	73.5	89.5	70.4	56.4	85.6	90.7	79.4	42.9	28.6	34.9	84.6	83.1	74.4	76.8
			Average TPM of expressed PRRs and R genes	11.2	11.7	12.3	8.2	6.4	10.5	13.8	3.0	19.1	3.4	7.1	5.0	8.4	9.2
	Citrus limon	SRR18218361	% of expressed transcripts	67.9	96.9	66.7	24.5	87.1	94.2	86.8	14.3	25.0	16.7	80.0	91.4	66.7	73.1
			Average TPM of expressed PRRs and R genes	3.0	3.3	31.2	5.3	2.1	5.0	2.9	0.1	3.1	7.3	2.6	4.3	6.4	5.9
	Citrus ichangensis	SRR10356230	% of expressed transcripts	66.7	87.0	68.2	41.7	83.3	88.9	73.0	28.8	22.2	20.0	100.0	100.0	59.1	58.9
			Average TPM of expressed PRRs and R genes	7.8	6.7	7.7	5.5	9.6	8.2	12.2	5.8	11.9	1.9	0.0	4.3	9.9	7.0
	Atalantia buxifolia	SRR13480641	% of expressed transcripts	70.0	82.1	63.0	45.8	86.8	84.1	82.5	38.9	40.0	52.9	100.0	84.6	68.8	70.5
			Average TPM of expressed PRRs and R genes	11.0	3.8	5.1	2.3	5.9	7.9	11.2	3.7	8.3	2.8	2.6	3.6	8.3	5.9

Table 3 Validation of the expression of clusters A and B of the RLK class in RNAseq libraries from leaves infected and uninfected by some type of pathogen

	Species	RNAseq libraries	RLK cluster	Total number of genes	N° of transcripts expressed	Average TPM of expressed transcripts
RNAseq libraries from leaves not infected by pathogens	*Citrus australis*	SRR22742835	A	8	8	18.13
			B	4	4	29.86
	Citrus reticulata	SRR25461105	A	9	9	8.05
			B	6	6	10.95
	Citrus medica	SRR15495635	A	14	12	1.74
			B1	3	1	2.04
			B2	5	5	2.96
	Citrus clementina	SRR7614644	A	11	11	5.47
	Citrus sinensis	SRR24101860	A	10	9	6.97
			B	4	4	5.98
	Citrus limon	SRR18218357	A	13	11	6.19
	Poncirus trifoliata	SRR7614646	A	17	16	4.53
	Fortunella hindsii	SRR9016287	A	15	15	11.29
	Species	RNAseq libraries	RLK cluster	Total number of genes	N° of transcripts expressed	Average TPM of expressed transcripts
RNAseq libraries of leaves infected by some type of pathogen	*Citrus maxima*	SRR23719598	A	9	9	7.68
			B	6	4	17.17
	Citrus reticulata	SRR14004272	A	9	8	8.90
			B	6	5	14.93
	Citrus medica	SRR14728280	A	14	14	5.05
			B1	3	3	15.78
			B2	5	5	4.41
	Citrus sinensis	SRR14004267	A	10	10	4.59
			B	4	4	27.63
	Citrus limon	SRR18218361	A	13	12	3.50

and Buddhas hand citron. All other *C. medica* accessions retained the full complement of cluster A genes.

Phylogenetic inference of RLK genes associated with resistance traits

Based on the phylogenetic analysis of RLK-class receptors, four orthologous groups were identified comprising proteins from citrus species resistant to citrus canker. The first group included sequences *from P. trifoliata, C. limon,* and *C. reticulata*; the second consisted of sequences from *C. limon, C. reticulata, F. hindsii*, and *C. micrantha*; while the third and fourth groups were formed by sequences from *F. hindsii, C. limon*, and *C. reticulata*. Two orthologous groups were associated with resistance to Alternaria brown spot, the first composed of sequences from *C. medica, C. clementina, C. limon,* and *C. micrantha*, and the second from *C. maxima, C. clementina, C. limon,* and *C. micrantha*. An additional orthologous group included sequences from species with reported resistance to both diseases, containing proteins from C. *maxima, C. medica*, and *C. limon*. One group was identified with sequences from P. trifoliata and C. reticulata, both species known to be resistant to *Phytophthora*. Finally, a conserved orthologous group containing sequences from all analyzed species wad also observed (Fig. 5).

Discussion

Reduced PRR and R gene content in wild citrus species compared to ancestral and hybrid lineages

Pattern recognition receptors (PRRs) and resistance (R) genes play a central role in plant immunity, acting in the recognition of PAMPsand/or effectors either at the cell surface or intracellularly, triggering defense responses with varying degrees of intensity depending on detection speed and specificity [18]. In this study, these genes were identified and compared across citrus species, revealing clear differences in their abundance. Wild species exhibited the lowest number of these genes, whereas ancestral species generally had higher counts, with hybrid descendants presenting intermediate profiles. These differences maybe primarily attributed to the varying levels of resistance to specific pathogens among species such as *C. sinensis*, *C. clementina*, *C. limon*, *C. australis*, *C. ichangensis*, *A. buxifolia*, *P. trifoliata*, and *F. hindsii* (Supplementary material 14), as well to distinct intensities of domestication and cultivation experienced by these lineages. Such heterogeneity likely imposed lineage-specific

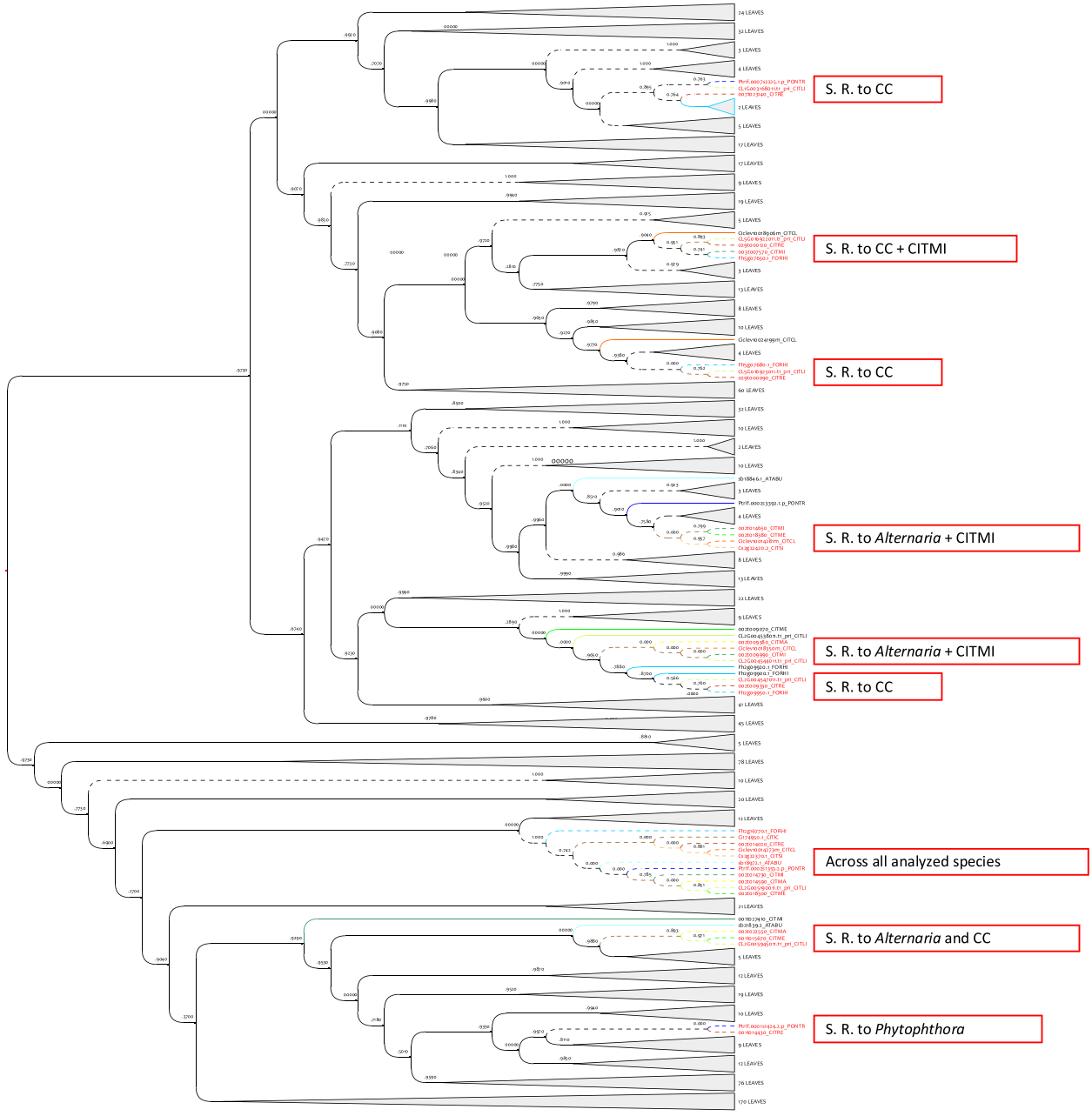

Fig. 5 Maximum likelihood tree based on RLK protein sequences from *Citrus reticulata* (CITRE), *C. maxima* (CITMA), *C. medica* (CITME), *C. micrantha* (CITMI), *C. sinensis* (CITSI), *C. clementina* (CITCL), *C. limon* (CITLI), *C. ichangensis* (CITIC), *A. buxifolia* (ATABU), *P. trifoliata* (PONTR), and *F. hindsii* (FORHI). Clades highlighted in red indicate orthologous gene groups identified in species resistant (S.R.) to citrus canker (CC), Alternaria brown spot (ABS), both diseases (CC + ABS), or *Phytophthora* spp. A conserved clade shared across all analyzed species is also shown

pressures that influenced gene retention, expansion, or loss over evolutionary time [45].

Gene expansion may have occurred through duplication events, which are among the primary mechanisms driving gene and genome evolution - providing raw material for neofunctionalization and diversification, particularly in plant immune systems [46–48]. Conversely, pseudogenization or loss of genes may result from reduced functional necessity or selective relaxation, supporting either the "less is more" hypothesis, or neutral evolutionary models such as "evolutionary regression" [49–51]. Furthermore, environmental factors such as temperature and humidity in citrus-growing regions likely influence pathogen pressure, and consequently, selection on immune receptor repertoires. Wild species, often subject to lower cultivation intensity and less consistent pathogen pressure, may retain only a minimal and functionally essential subset of immune-related genes.

Similar patterns of reduced immune gene content in wild versus cultivated species have been reported in soybean (*Glycine max* vs. *Glycine latifolia*) [52], tomato (*Solanum lycopersicum* vs. *Solanum pimpinellifolium*) [53], wheat (*Triticum aestivum* vs. *Aegilops tauschii* and *Triticum dicoccoides*) [54], and rice (*Oryza sativa* spp. vs. *Oryza nivara*, and *O. sativa* Kalasath vs. *O. nivara*, *Oryza rufipogon*, and *Oryza barthii*) [20, 55]. These findings are often attributed to high metabolic costs of maintaining immune receptors in the absence of pathogens, leading wild species to conserve a limited yet sufficient repertoire of defense genes. Moreover, while wild individuals may harbor fewer resistance genes individually, their collective genetic diversity can provide a rich source of novel immune alleles across populations [56].

Predominance of membrane-localized PRR receptors in wild citrus species

Among the classes of receptors analyzed, RLK, CNL, and TNL proteins stands out due their central role in plant immune signaling. Receptor-like kinases (RLKs), which are localized at the plasma membrane, are primarily responsible for the recognition of PAMPs and initiating PAMP-triggered immunity (PTI), the first layer of plant defense [15, 57, 58]. In this study, RLK genes represented a notable proportion of immune receptor repertoire, especially in *C. australis* and *A. buxifolia*, where they were the most abundant among PRR and R genes classes. In addition, elevated proportions of RLP and RLKGNK2 receptors were observed in *C. australis* and *C. ichangensis*.

These findings reveal a predominance of membrane-localized PRR receptors in wild citrus species, which may reflect a functional emphasis on early detection of extracellular pathogens. However, further experimental studies are needed to validate this functional implication. PTI responses, while generally less specific than effector-triggered immunity (ETI), incur lower metabolic costs and are effective at containing a broad spectrum of pathogens. Because PTI activation proceeds more gradually, it allows the plant to assess whether the initial responses are sufficient. If successful, downstream signaling is attenuated to avoid unnecessary energy expenditure [18].

Furthermore, wild citrus species serve as valuable reservoirs of immune-related genes, which are routinely incorporated into breeding programs to enhance disease resistance in cultivated varieties. This strategy has been widely adopted in other crops such as cassava, wheat, millet, rice, corn, sunflower, lettuce, banana, potato, peanut, and tomato [59, 60], underlining the agronomic of maintaining and characterizing wild germplasm.

Ancestral and hybrid citrus species exhibit predominance of CNL and TNL receptors

CNL and TNL proteins are intracellular immune receptors that detect pathogen-derived effectors and activate effector-triggered immunity (ETI), the second layer of plant defense [15, 61]. In this study, a predominance of CNL receptors were observed across the four ancestral species, as well as in *C. sinensis*, *C. clementina*, *C. limon*, *P. trifoliata*, and particularly in *F. hindsii*. Similarly, TNL receptors were a notably represented in these species, with exception of *C. maxima*.

These two classes share conserved nucleotide-binding (NB or NBS) and leucine-rich repeat (LRR) domains but differ at the N-terminal region: CNL proteins possess a coiled-coil domain, whereas TNL proteins contain a Toll/interleukin-1 receptor (TIR) domain [36, 62]. ETI responses mediated by these receptors are generally more specific and robust than PTI, although they require greater metabolic investment and subject to modulation by developmental environmental cues [18].

The prevalence of CNL and TNL classes in both ancestral and hybrid species suggests an evolutionary expansion of intracellular immune receptors, likely driven by the necessity to recognize a broader repertoire of effectors in cultivated environments. Similar patterns have been reported in other species, including *Medicago truncatula*, *Cajanus cajan*, *Phaseolus vulgaris*, *G. max*, *Vitis vinifera*, *Capsicum annuum*, *S. lycopersicum*, and *S. tuberosum*, where CNL genes consistently represent the most abundant class among resistance genes [63].

Furthermore, wild citrus species in this study exhibited a markedly reduced number of TNL-type proteins, consistent with findings in other plant systems. For example, in *G. latifolia* a wild relative of soybean, the number of TNL genes was approximately four times lower than in the cultivated *G. max* genome [52, 64]. These results reinforce the idea that artificial selection and domestication processes may have contributed to the expansion of ETI-related receptors in cultivated species, while wild species tend to conserve a smaller but essential subset of these genes.

Phylogenetic analysis correlating PRR and R gene classes with potential pathogen response specificity have been conducted in a complementary study (manuscript in preparation), supporting the observed variation between citrus species and suggesting a possible co-evolutionary trajectory.

Citrus australis exhibits enhanced basal expression of PRR and R genes in healthy leaf tissues

Analysis of RNA-seq libraries from healthy leaf tissues revealed that *C. australis* presented the highest average TPM for multiple PRR and R gene classes, despite possessing a comparatively lower total number of identified

receptors. Notably, elevated expression levels were observed for plasma membrane-localized receptors (RLK and RLP) as well as cytoplasmic classes (CNL, N, NL, and RPW8NL), all exceeding the average TPM for the respective library.

This expression pattern suggests a transcriptional profile consistent with a wide activation of immune-related genes in *C. australis*, relative to the average TPM values within its own RNA-seq dataset, though no cross-species comparisons were made due to differences in experimental conditions. Such a profile may indicate a more active basal defense system in *C. australis*, potentially involving both PTI and pre-formed or constitutive defense mechanisms. Nonetheless, functional validation is required to determine whether this expression pattern translates into enhanced disease resistance.

Basal resistance is characterized as the first line of defense that restricts pathogen growth following initial infection, and is typically mediated by rapid production of reactive oxygen species, activation of mitogen-activated protein kinase cascades, transcriptional reprogramming of defense-related genes [65, 66].

In citrus, basal resistance mechanisms have been associated with tolerance to HLB a major bacterial disease affecting citrus crops worldwide [67]. Notably, *C. australis* is among the few species reported to exhibit natural resistance to HLB [32]. Therefore, the elevated expression of both membrane-bound and intracellular immune receptors in the absence of pathogen pressure may reflect a constitutive state of preparedness in *C. australis*; however, whether this correlates with increased disease tolerance remains to be experimentally confirmed.

Citrus reticulata and *C. medica* exhibit broad transcriptional activation of PRR RNAs and R genes during pathogen infection

In the RNA-seq libraries derived pathogen-infected citrus leaves, *C. medica* and *C. reticulata* demonstrated expression levels above the species-specific average across a greater number of PRR and R gene classes when compared to other species. This broader transcriptional activation may reflect the plants' response to effector molecules secreted by the invading pathogens, likely triggering elevated both membrane-associated and cytoplasmic immune receptors in an effort to mount a more robust defense response [15, 18].

The upregulation of diverse immune gene classes in these species is consistent with their known phenotypic resistance to multiple diseases. *C. reticulata* has been reported as resistant to citrus gummosis, citrus canker, and citrus variegated chlorosis [68–70], while *C. medica* exhibits resistance to Alternaria brown spot and citrus canker [71, 72]. These transcriptional patterns may reflect a more responsive and possibly more efficient immune signaling network in these species when exposed to biotic stress; however, further studies are required to evaluate the effectiveness of these responses against specific pathogens.

PRR receptors and R genes are predominantly located on chromosomes 1, 2, 3, 5, and 7

The chromosomal mapping of PRR and R genes across citrus genomes revealed that these immune receptors are predominantly distributed on chromosomes 1, 2, 3, 5, and 7. Specifically, genes from the CNL class were more frequently located on chromosomes 5 and 7; TNL genes were enriched on chromosome 3, and RLK genes were predominantly distributed on chromosomes 1 and 2. These findings align with prior molecular mapping studies that identified quantitative trait loci (QTLs) associated with disease resistance on the same chromosomes.

In *P. trifoliata*, a species tolerant to HLB and resistant to *Phytophthora*, QTLs were mapped to Linkage groups 2, 3, 5, and 7 in response to *Candidatus* Liberibacter asiaticus infection [73, 74], and to groups 3, 5, and 7 following *Phytophthora* challenge [75]. Similarly, in *Citrus sunki*, a susceptible species, QTLs were reported on Linkage groups 2, 3, 5, and 7 under similar infection conditions [73, 75]. In hybrids derived from Clementina x Tangelo, QTLs related to resistance against *Citrus chlorotic dwarf-associated virus* were detected on Linkage groups 2, 3, and 5 [76].

Furthermore, in *Poncirus polyandra*, immune receptors were concentrated on chromosomes 1, 2, and 3, with HLB tolerance loci identified on chromosomes 2 and 3 and resistance to *Closterovirus tristezae* (CTV) mapped to chromosomes 1, 5, and 7 [77]. This distribution pattern closely mirrors the chromosomal localization of RLK genes observed in *C. australis* in the present study. Together, these data suggest a functional clustering of resistance-related loci that may reflect evolutionary conservation of key defense pathways in citrus.

All analyzed citrus species exhibit duplication events in the RLK receptor class

Gene duplication is a major force evolutionary contributing to plant adaptability, particularly under biotic and abiotic stress conditions. In plants, duplications events may occur through various mechanisms, including tandem duplication (adjacent genes copies), segmental or whole-genome duplication (WGD), dispersed duplication (copies spread across distant genomic regions), and proximal duplication (copies located nearby but not adjacent). These events contribute to gene family expansion and provide raw material for evolutionary innovation. The retention of duplicated PRR and R genes is often shaped by purifying selection, which preserves essential functions, or by positive selection, which favors

functional divergence. Over time, duplicated genes may undergo subfunctionalization (partitioning of ancestral functions) or neofunctionalization (acquisition of novel functions), both of which can enhance immune versatility and response diversity [45, 78, 79].

In the present study, all evaluated citrus species exhibited gene duplication events within the RLK receptor class. Notably, *C. australis*, *P. trifoliata*, and *F. hindsii* displayed the most diversified patterns, with duplicated genes distributed across multiple chromosomes (chromosomes 1, 2, 4, and 7), suggesting enhanced genomic plasticity in these species [80].

This widespread duplication of RLK genes may reflect a selective advantage by expanding the repertoire of plasma membrane-localized receptors involved in the recognition of PAMPs, thereby potentially enhancing the efficiency and speed of PTI in these species., Moreover duplicated RLK genes may diverge functionally over time contributing to the evolution of novel immune responses [47]. For example, in *S. lycopersicum*, a duplicated RLK gene to have acquired a function distinct from its ancestral counterpart, representing a clear adaptative gain in pathogen recognition and defense [81].

These observations underscore the importance of gene duplication in shaping immune diversity in citrus and support the hypothesis that receptor expansion through duplication events may play a critical role in the coevolutionary dynamics between citrus and its pathogens.

PRR receptors and R genes in citrus are predominantly organized in genomic clusters

Genomic clustering is a common organizational feature of PRR receptors and R genes in plant genomes, resulting from evolutionary processes such as tandem duplication and transposable element activity [19, 82]. In this study, such clustered was observed across all analyzed species, with variation in clustering levels among species. *C. australis*, in particular, exhibited a comparatively lower percentage of clustered genes, suggesting reduced local duplication activity or distinct evolutionary trajectories.

Clusters were predominantly located on chromosomes 2, 3, 5, or 7 across species. These chromosomal regions are consistent with known hotspots for resistance loci in citrus and other plant genomes [19, 26, 27, 82–92]. Among RLK gene clusters selected for detailed analysis, full expression of all cluster-associated transcripts was observed in *C. australis* and *C. reticulata* under non-infected conditions, and *C. medica* and *C. sinensis* under pathogen infection. In *C. australis* and *C. reticulata*, these findings are consistent with the "less is more" hypothesis of evolutionary gene retention, whereby gene loss over time may favor the maintenance of a smaller set of functionally essential genes with stronger or more efficient responses; although confirmation through functional studies is needed [49, 51]. The high expression of RLK clusters in these species suggest a prominent role for PTI in their basal defense mechanisms, consistent with previously reported HLB tolerance in *C. australis* [67], and for resistance to variegated chlorosis in *C. reticulata* [93].

Conversely, the full cluster expression observed in *C. medica* under infected conditions may reflect enhanced transcriptional activation following effector perception, potentially linked to previously reported resistance to *Alternaria alternata* and *Xanthomonas citri* [71, 72]. In *C. sinensis*, despite high expression levels under infection, its known susceptibility to several citrus pathogens [68, 69, 75, 77, 94–97] suggests that these responses may not be sufficient to restrict disease progression, possibly due to the inefficiency or late activation of defense signaling pathways.

Citrus species exhibit intraspecific diversity of PRR and R genes

Citrus species are known for their high levels of genetic diversity [2, 98–100], yet comprehensive studies addressing intraspecific variation in PRR and R gene content remain limited. Our analysis of four ancestral species revealed considerable intraspecific variation in gene copy number, particularly for CNL class in *C. micrantha* and *C. reticulata*. The RLK class also exhibited notable variability within these species. In addition, intraspecific differences in the gene content of RLK clusters were detected in *C. reticulata* and *C. medica*. Such diversity is essential for breeding programs, as it enables the identification of superior genotypes that can be incorporated into hybridization strategies or targeted for disease-resistance improvement using molecular tools [101–103].

Similar patterns of intraspecific variation in PRR and R genes have been reported in *Arabidopsis thaliana* [104, 105], *Cucumis melo* [106], *O. sativa* [107, 108], and *Triticum dicoccoides* [109], reinforcing the significance of this diversity for adaptative immunity and crop improvement (Table 4).

Conclusion

PRR receptors and R genes in citrus are predominantly distributed across chromosomes 1, 2, 3, 5, and 7, with a notable representation of the CNL, NL, RLK, and TNL classes. These gene families exhibit both interspecific and intraspecific diversity, with the CNL class showing the highest levels of variation, while MLO and RPW8NL remain largely conserved. Wild citrus species, including *C. australis*, *C. ichangensis*, and *A. buxifolia*, possess fewer PRR and R genes overall - particularly those with TIR domains - compared to both ancestral and hybrid species.

Table 4 Intra-specific comparison for the quantity of PRR receptors and R genes belonging to the classes CN, CNL, MLO, N, NL, RLK, RLKGNK2, RLP, RPW8NL, T, TN, and TNL between *C. maxima* and 13 citron accessions, *C. reticulata* and 16 mandarin accessions , *C. micrantha* and three papeda accessions, and *C. medica* and 11 citron accessions

		CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	TOTAL
PAV analysis of CITMA	CITMA	41	227	21	48	104	95	32	18	7	14	5	19	631
	pum_hf22D	34	184	21	42	88	94	30	17	7	13	4	18	552
	pum_GuilD	34	187	19	41	86	92	29	15	7	10	4	17	541
	pum_tahiD	34	175	18	40	83	92	28	16	7	13	4	17	527
	pum_cha2D	33	157	19	41	79	92	29	15	7	14	5	17	508
	pum_HuazD	31	164	17	39	79	91	29	15	7	12	4	16	504
	pum_HuanD	29	154	19	39	75	93	26	15	7	14	5	18	494
	pum_timoD	24	154	18	38	81	91	31	16	7	12	4	17	493
	pum_reinD	32	156	21	35	70	93	27	15	7	13	4	19	492
	pum_kaopD	31	143	19	38	77	90	28	14	7	14	4	18	483
	pum_florD	30	150	18	37	72	92	28	16	7	12	3	17	482
	pum_indeD	28	135	18	39	76	90	28	17	7	13	5	17	473
	pum_ChonD	32	143	21	33	65	92	27	16	7	12	4	13	465
	pum_chaH	31	129	17	37	65	87	27	15	7	11	4	15	445
	Standard Deviation	3,81	25,38	1,47	3,52	10,09	1,94	1,65	1,07	0,00	1,22	0,58	1,57	46,01
	CV (%)	12,0	15,7	7,7	9,0	12,8	2,1	5,8	6,8	0,0	9,6	13,7	9,2	9,1

		CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	TOTAL
PAV analysis of CITRE	CITRE	37	250	17	44	147	102	30	26	12	19	11	56	751
	man_clo1D	35	248	17	40	143	101	30	24	12	19	10	56	735
	man_ponkD	32	187	17	30	107	96	26	22	11	15	9	39	591
	man_shekD	29	179	17	35	104	95	29	21	12	12	7	40	580
	man_DaoW1D	25	187	17	28	104	89	27	20	12	13	2	42	566
	man_ManWD	25	185	17	33	109	82	27	17	12	12	4	40	563
	man_HezWD	24	176	17	31	103	83	28	22	12	14	10	41	561
	man_scshD	27	172	17	28	106	84	26	23	12	15	9	42	561
	man_szinD	25	186	17	29	101	89	27	19	12	13	2	41	561
	man_JiaWD	27	176	17	30	105	80	28	18	12	14	7	43	557
	man_kingD	23	185	17	31	97	92	26	20	12	12	3	37	555
	man_daoxD	28	176	17	29	104	80	27	18	12	14	7	42	554
	man_nfmcD	25	168	17	30	103	92	29	17	12	12	9	33	547
	man_tachD	25	166	17	26	93	93	28	19	12	15	10	42	546
	man_sunkD	25	168	17	36	94	84	27	21	12	12	10	39	545
	man_owarD	22	175	17	27	93	89	26	20	12	11	4	36	532
	man_avanD	27	167	17	27	89	85	26	20	12	12	3	37	522
	Standard Deviation	4,09	25,01	0,00	4,86	15,75	6,82	1,37	2,45	0,24	2,33	3,20	6,04	64,16
	CV (%)	15,1	13,5	0,0	15,5	14,9	7,6	5,0	12,0	2,0	16,9	46,5	14,6	11,1

		CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	TOTAL
PAV analysis of CITMI	CITMI	20	182	17	52	100	97	35	17	6	35	26	73	660
	pap_mic1D	17	176	17	47	95	97	35	17	6	31	25	67	630
	pap_macrD	10	140	17	39	73	95	28	12	6	26	17	54	517
	pap_melaD	12	117	17	20	59	84	25	13	6	18	12	41	424
	Standard Deviation	4,57	30,73	0,00	14,06	19,17	6,24	5,06	2,63	0,00	7,33	6,68	14,24	108,36
	CV (%)	31,0	20,0	0,0	35,6	23,5	6,7	16,4	17,8	0,0	26,6	33,4	24,2	19,4

		CN	CNL	MLO	N	NL	RLK	RLKGNK2	RLP	RPW8NL	T	TN	TNL	TOTAL
PAV analysis of CITME	CITME	37	139	18	54	83	88	36	25	14	33	15	55	597
	cit_cor1D	36	137	18	53	82	88	36	25	14	32	14	55	590
	cit_diamD	34	137	18	52	82	88	34	25	14	32	15	55	586
	cit_wc1D	32	134	18	50	81	88	36	23	14	30	14	51	571
	cit_BdHD	33	134	18	51	79	87	35	24	14	30	12	51	568
	cit_wc2D	34	130	18	50	82	87	35	24	14	27	13	51	565
	cit_macvD	33	133	18	50	80	86	35	24	14	28	13	48	562
	cit_e861D	31	132	18	52	75	87	32	25	13	31	14	50	560
	cit_chinD	33	131	18	49	75	88	34	24	14	26	13	48	553
	cit_wc3D	34	131	18	47	74	87	36	24	14	27	13	46	551
	cit_humpD	30	132	18	50	75	86	32	24	14	26	13	48	548
	cit_macnD	30	132	18	48	78	87	34	24	14	25	12	45	547
	Standard Deviation	2,15	2,81	0,00	2,02	3,33	0,75	1,44	0,62	0,29	2,75	1,00	3,44	16,77
	CV (%)	6,5	2,1	0,0	4,0	4,2	0,9	4,2	2,6	2,1	9,5	7,4	6,8	3,0

*The color scale varies in each column, from green to red, according to the highest similarity in the quantity of identified receptors in a specific access compared to the ancestral species used as reference. CV (%) represents the coefficient of variation in percentage

A marked predominance of membrane-associated PRR receptors, including RLKs and RLPs, was observed in wild species, which could be associated with early mechanisms. In contrast, intracellular receptors from the CNL and TNL classes were more prominent in ancestral and hybrid species. These patterns may reflect distinct immune strategies, but further functional studies are needed to confirm their roles in PTI or ETI responses.

Gene duplication and clustering also played key roles in shaping PRR/R gene repertoires, with RLK gene clusters identified in all species and showing functional expression in both healthy and pathogen-infected conditions. *C. australis* and *C. reticulata* displayed consistent expression of RLK clusters under basal conditions, while C. *medica* and *C. sinensis* showed cluster-wide expression during infection. Intraspecific variation in gene number and cluster composition was especially pronounced in *C. micrantha* and *C. reticulata*, reinforcing their potential value in breeding programs.

Together, these findings highlight the structural, functional, and evolutionary diversity of immune receptor genes across citrus species and support their relevance as molecular resources for disease resistance breeding and genomic selection.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-12043-5.

Supplementary Material 1

Acknowledgements
FSM was funded by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). RJSS was funded by CAPES. FM received a Productivity Grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Authors' contributions
FSM was responsible for the execution of the all the experimental steps; RJSS participated to PRR and R genes identification; RJSS, PO, and FM were responsible for the conception and design of the experiments; FSM and FM wrote the manuscript; FM and PO were responsible for the financial support of the research; FM, RJSS and PO advised FSM.

Funding
This study received funds from Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

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

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 6 November 2024 / Accepted: 25 August 2025
Published online: 14 January 2026

References

1. Luro F, Curk F, Froelicher Y, Ollitrault P: Recent Insights on Citrus Diversity and Phylogeny. In: AGRUMED: Archaeology and History of Citrus Fruit in the Mediterranean. Edited by Zech-Matterne V, Fiorentino G: Publications du Centre Jean Bérard; 2017.
2. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of *Citrus*. Nature. 2018;554:311.
3. Liu Y, Heying E, Tanumihardjo SA. History, global distribution, and nutritional importance of citrus fruits. Compr Rev Food Sci Food Saf. 2012;11(6):530–45.
4. IBGE: Instituto Brasileiro de Geografia e Estatistica. https://www.ibge.gov.br/. Accessed in 2023.
5. USDA: Citrus: World Markets and Trade. https://esmis.nal.usda.gov/publication/citrus-world-markets-and-trade. Accessed in 2023.
6. Zhong G, Nicolosi E. Citrus Origin, Diffusion, and Economic Importance. In: Gentile A, La Malfa S, Deng Z, editors. The Citrus Genome. Cham: Springer International Publishing; 2020. p. 5–21.
7. Bassanezi RB, Lopes SA, de Miranda MP, Wulff NA, Volpe HXL, Ayres AJ. Overview of citrus huanglongbing spread and management strategies in Brazil. Trop Plant Pathol. 2020;45(3):251–64.
8. Behlau F. An overview of citrus canker in Brazil. Trop Plant Pathol. 2021;46(1):1–12.
9. Coletta-Filho HD, Castillo AI, Laranjeira FF, de Andrade EC, Silva NT, de Souza AA, et al. Citrus Variegated Chlorosis: an Overview of 30 Years of Research and Disease Management. Tropical Plant Pathology. 2020;45(3):175–91.
10. Folimonova SY, Sun Y-D. Citrus Tristeza Virus: From Pathogen to Panacea. Annu Rev Virol. 2022;9(9):417–35.
11. Chabi-Jesus C, Ramos-González PL, Postclam-Barro M, Fontenele RS, Harakava R, Bassanezi RB, et al. Molecular epidemiology of Citrus leprosis virus C: a new viral lineage and phylodynamic of the main viral subpopulations in the Americas. Front Microbiol. 2021. https://doi.org/10.3389/fmicb.2021.641252.
12. Gaikwad PN, Sharma V, Singh J, Sidhu GS, Singh H, Omar AA. Biotechnological advancements in *Phytophthora* disease diagnosis, interaction and management in citrus. Sci Hortic. 2023;310:111739.
13. Santos Dória M, Silva Guedes M, de Andrade Silva EM. Magalhães de Oliveira T, Pirovani CP, Kupper KC, Bastianel M, Micheli F: Comparative proteomics of two citrus varieties in response to infection by the fungus Alternaria alternata. Int J Biol Macromol. 2019;136:410–23.
14. Agrios G: Plant Pathology, 5th Edition edn: Academic Press; 2005:922.
15. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444(7117):323–9.
16. Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. Science (New York, NY). 2013;341(6147):746–51.
17. Boller T, He SY. Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. Science. 2009;324(5928):742–4.
18. Katagiri F, Tsuda K. Understanding the plant immune system. Mol Plant Microbe Interact. 2010;23(12):1531–6.
19. Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. 1998;8(11):1113–30.
20. Mizuno H, Katagiri S, Kanamori H, Mukai Y, Sasaki T, Matsumoto T, et al. Evolutionary dynamics and impacts of chromosome regions carrying R-gene clusters in rice. Sci Rep. 2020;10(1):872.
21. Bergelson J, Kreitman M, Stahl EA, Tian D. Evolutionary dynamics of plant R-genes. Science. 2001;292(5525):2281–5.
22. McHale L, Tan X, Koehl P, Michelmore RW. Plant NBS-LRR proteins: adaptable guards. Genome Biol. 2006;7(4):212.
23. Baumgarten A, Cannon S, Spangler R, May G. Genome-level evolution of resistance genes in *Arabidopsis thaliana*. Genetics. 2003;165(1):309–19.
24. Chovelon V, Feriche-Linares R, Barreau G, Chadoeuf J, Callot C, Gautier V, et al. Building a cluster of NLR genes conferring resistance to pests and pathogens: the story of the Vat gene cluster in cucurbits. Hortic Res. 2021;8(1):72.
25. Hulbert SH, Webb CA, Smith SM, Sun Q. Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol. 2001;39:285–312.

26. van Wersch S, Li X. Stronger when together: clustering of plant NLR disease resistance genes. Trends Plant Sci. 2019;24(8):688–99.
27. Wang Y, Zhou L, Li D, Dai L, Lawton-Rauh A, Srimani PK, et al. Genome-wide comparative analysis reveals similar types of NBS genes in hybrid *Citrus sinensis* genome and original *Citrus clementine* genome and provides new insights into non-TIR NBS genes. PLoS ONE. 2015;10(3):e0121893.
28. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al. The draft genome of sweet orange (Citrus sinensis). Nat Genet. 2012;45:59.
29. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotechnol. 2014;32(7):656–62.
30. Guardo MD, Moretto M, Moser M, Catalano C, Troggio M, Deng Z, et al. The haplotype-resolved reference genome of lemon (*Citrus limon* L. Burm f.). Tree Genet Genomes. 2021;17(6):46.
31. Wang X, Xu Y, Zhang S, Cao L, Huang Y, Cheng J, et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. Nat Genet. 2017;49(5):765–72.
32. Nakandala U, Masouleh AK, Smith MW, Furtado A, Mason P, Constantin L, et al. Haplotype resolved chromosome level genome assembly of *Citrus australis* reveals disease resistance and other citrus specific genes. Hortic Res. 2023. https://doi.org/10.1093/hr/uhad058.
33. Peng Z, Bredeson JV, Wu GA, Shu S, Rawat N, Du D, et al. A chromosome-scale reference genome of trifoliate orange (*Poncirus trifoliata*) provides insights into disease resistance, cold tolerance and genome evolution in Citrus. Plant J. 2020;104(5):1215–32.
34. Wang N, Song X, Ye J, Zhang S, Cao Z, Zhu C, et al. Structural variation and parallel evolution of apomixis in citrus during domestication and diversification. Natl Sci Rev. 2022;9(10):nwac114.
35. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. Interproscan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40.
36. Santana Silva RJ, Micheli F. Rrgpredictor, a set-theory-based tool for predicting pathogen-associated molecular pattern receptors (PRRs) and resistance (R) proteins from plants. Genomics. 2020;112(3):2666–76.
37. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
38. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCscanx: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–e49.
39. Jupe F, Pritchard L, Etherington GJ, Mackenzie K, Cock PJ, Wright F, et al. Identification and localisation of the NB-LRR gene family within the potato genome. BMC Genomics. 2012;13:75.
40. Holub EB. The arms race is ancient history in *Arabidopsis*, the wildflower. Nat Rev Genet. 2001;2(7):516–27.
41. Tay Fernandez CG, Marsh JI, Nestor BJ, Gill M, Golicz AA, Bayer PE, et al. An SGSGeneloss-based method for constructing a gene presence-absence table using Mosdepth. Methods Mol Biol. 2022;2512:73–80.
42. Katoh K, Standley DM. MAFFT multiplesequence alignment software version 7: improvements in performance and usability. Mol BiolEvol. 2013;30(4):772–80. https://doi.org/10.1093/molbev/mst010. Epub 2013 Jan 16. PMID:23329690; PMCID: PMC3603318.
43. Capella-Gutiéérrez S, Silla-Martíinez JM, Gabaldóón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
44. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0, Systematic Biology. 2010;59:307–21.
45. Gonzalez-Ibeas D, Ibanez V, Perez-Roman E, Borredá C, Terol J, Talon M. Shaping the biology of citrus: I. Genomic determinants of evolution. Plant Genome. 2021;14(3):e20104.
46. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. J Genet. 2013;92(1):155–61.
47. Panchy N, Lehti-Shiu M, Shiu S-H. Evolution of gene duplication in plants. Plant Physiol. 2016;171(4):2294–316.
48. Roth C, Rastogi S, Arvestad L, Dittmar K, Light S, Ekman D, et al. Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. J Exp Zool B Mol Dev Evol. 2007;308(1):58–73.
49. Albalat R, Cañestro C. Evolution by gene loss. Nat Rev Genet. 2016;17(7):379–91.
50. Clark JW. Genome evolution in plants and the origins of innovation. New Phytol. 2023;240(6):2204–9.
51. Olson MV. When less is more: gene loss as an engine of evolutionary change. Am J Hum Genet. 1999;64(1):18–23.
52. Liu Q, Chang S, Hartman GL, Domier LL. Assembly and annotation of a draft genome sequence for *Glycine latifolia*, a perennial wild relative of soybean. Plant J. 2018;95(1):71–85.
53. Wei H, Liu J, Guo Q, Pan L, Chai S, Cheng Y, et al. Genomic organization and comparative phylogenic analysis of NBS-LRR resistance gene family in *Solanum pimpinellifolium* and *Arabidopsis thaliana*. Evol Bioinform Online. 2020;16:1176934320911055.
54. Ye C-Y, Wu D, Mao L, Jia L, Qiu J, Lao S, et al. The genomes of the allohexaploid *Echinochloa crus-galli* and its progenitors provide insights into polyploidization-driven adaptation. Mol Plant. 2020;13(9):1298–310.
55. Stein JC, Yu Y, Copetti D, Zwickl DJ, Zhang L, Zhang C, et al. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. Nat Genet. 2018;50(2):285–96.
56. Barabaschi D, Tondelli A, Valè G, Cattivelli L. Fitness cost shapes differential evolutionary dynamics of disease resistance genes in cultivated and wild plants. Mol Plant. 2020;13(10):1352–4.
57. He K, Wu Y. Receptor-like kinases and regulation of plant innate immunity. Enzymes. 2016;40:105–42.
58. Liu P-L, Du L, Huang Y, Gao S-M, Yu M. Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol Biol. 2017;17(1):47.
59. Okoń S, Ociepa T, Nucia A, Cieplak M, Kowalczyk K. Is every wild species a rich source of disease resistance? *Avena fatua* L.—potential donor of resistance to powdery mildew. Plants. 2021;10(3):560.
60. Hajjar R, Hodgkin T. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica. 2007;156(1):1–13.
61. Contreras MP, Lüdke D, Pai H, Toghani A, Kamoun S. NLR receptors in plant immunity: making sense of the alphabet soup. EMBO Rep. 2023;24(10):e57495.
62. Takken FL, Goverse A. How to build a pathogen detector: structural basis of NB-LRR function. Curr Opin Plant Biol. 2012;15(4):375–84.
63. Shao ZQ, Xue JY, Wu P, Zhang YM, Wu Y, Hang YY, et al. Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. Plant Physiol. 2016;170(4):2095–109.
64. Kang YJ, Kim KH, Shim S, Yoon MY, Sun S, Kim MY, et al. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biol. 2012;12(1):139.
65. Boller T, Felix G. A renaissance of elicitors: perception of Microbe-Associated Molecular Patterns and danger signals by Pattern-Recognition Receptors. Annu Rev Plant Biol. 2009;60(1):379–406.
66. Dangl JL, Jones JD. Plant pathogens and integrated defence responses to infection. Nature. 2001;411(6839):826–33.
67. Wang Y, Zhou L, Yu X, Stover E, Luo F, Duan Y. Transcriptome Profiling of Huanglongbing (HLB) Tolerant and Susceptible Citrus Plants Reveals the Role of Basal Resistance in HLB Tolerance. Front Plant Sci 2016, 7.
68. de Carvalho SA, de Carvalho Nunes WM, Belasque J Jr, Machado MA, Croce-Filho J, Bock CH, et al. Comparison of Resistance to Asiatic Citrus Canker Among Different Genotypes of Citrus in a Long-Term Canker-Resistance Field Screening Experiment in Brazil. Plant Dis. 2015;99(2):207–18.
69. Niza B, Coletta-Filho HD, Merfa MV, Takita MA, de Souza AA. Differential colonization patterns of *Xylella fastidiosa* infecting citrus genotypes. Plant Pathol. 2015;64(6):1259–69.
70. Yan H, Zhong Y, Jiang B, Zhou B, Wu B, Zhong G. Guanggan (Citrus reticulata) shows strong resistance to Phytophthora nicotianae. Sci Hortic. 2017;225:141–9.
71. Deng ZN, Xu L, Li DZ, Long GY, Liu LP, Fang F, et al. Screening citrus genotypes for resistance to canker disease (*Xanthomonas axonopodis* pv. *citri*). Plant Breed. 2010;129(3):341–5.
72. Solel Z, Kimchi M. Susceptibility and resistance of Citrus genotypes to *Alternaria alternata* pv. citri. J Phytopathol. 1997;145(8–9):389–91.
73. Soratto TAT, Curtolo M, Marengo S, Dezotti AL, Lima RPM, Gazaffi R, et al. QTL and eQTL mapping associated with host response to *Candidatus* Liberibacter asiaticus in citrandarins. Trop Plant Pathol. 2020;45(6):626–45.
74. Huang M, Roose ML, Yu Q, Stover E, Hall DG, Deng Z, et al. Mapping of QTLs and candidate genes associated with multiple phenotypic traits for Huanglongbing tolerance in citrus. Hortic Plant J. 2023;9(4):705–19.
75. Lima RPM, Curtolo M, Merfa MV, Cristofani-Yaly M, Machado MA. QTLs and eQTLs mapping related to citrandarins' resistance to citrus gummosis disease. BMC Genomics. 2018;19(1):516.

76. Aksoy D, Yeşiloğlu T, Serçe ÇU, Ayaz F. Determination of genomic regions related citrus chlorotic dwarf disease in clementine mandarin × Minneola tangelo hybrid population by QTL analysis. Euphytica. 2023;219(2):24.
77. Zhang S, Chen J, Zhang C, Zhang S, Zhang X, Gao L, et al. Insights into identifying resistance genes for cold and disease stresses through chromosome-level reference genome analyses of *Poncirus polyandra*. Genomics. 2023;115(3):110617.
78. Kondrashov FA. Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc Biol Sci. 2012;279(1749):5048–57.
79. Qian W, Zhang J. Genomic evidence for adaptation by gene duplication. Genome Res. 2014;24(8):1356–62.
80. Zhang J. Evolution by gene duplication: an update. Trends Ecol Evol. 2003;18(6):292–8.
81. Buendia L, Wang T, Girardin A, Lefebvre B. The LysM receptor-like kinase SILYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. New Phytol. 2016;210(1):184–95.
82. Lee RRQ, Chae E. Variation patterns of NLR clusters in *Arabidopsis thaliana* genomes. Plant Commun. 2020;1(4):100089.
83. Van de Weyer AL, Monteiro F, Furzer OJ, Nishimura MT, Cevik V, Witek K, et al. A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis thaliana. Cell. 2019;178(5):1260-1272.e1214.
84. Richly E, Kurth J, Leister D. Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. Mol Biol Evol. 2002;19(1):76–84.
85. Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, et al. Genome-wide identification of NBS genes in *japonica* rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Mol Genet Genomics. 2004;271(4):402–15.
86. Ameline-Torregrosa C, Wang BB, O'Bleness MS, Deshpande S, Zhu H, Roe B, et al. Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. Plant Physiol. 2008;146(1):5–21.
87. Kohler A, Rinaldi C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, et al. Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. Plant Mol Biol. 2008;66(6):619–36.
88. Arya P, Kumar G, Acharya V, Singh AK. Genome-wide identification and expression analysis of NBS-encoding genes in *Malus x domestica* and expansion of NBS genes family in Rosaceae. PLoS ONE. 2014;9(9):e107987.
89. Lozano R, Hamblin MT, Prochnik S, Jannink JL. Identification and distribution of the NBS-LRR gene family in the Cassava genome. BMC Genomics. 2015;16(1):360.
90. Seo E, Kim S, Yeom S-I, Choi D. Genome-Wide Comparative Analyses Reveal the Dynamic Evolution of Nucleotide-Binding Leucine-Rich Repeat Gene Family among Solanaceae Plants. Front Plant Sci 2016;7.
91. Li X, Salman A, Guo C, Yu J, Cao S, Gao X, Li W, Li H, Guo Y. Identification and Characterization of LRR-RLK Family Genes in Potato Reveal Their Involvement in Peptide Signaling of Cell Fate Decisions and Biotic/Abiotic Stress Responses. Cells 2018,7(9).
92. Yan J, Su P, Meng X, Liu P. Phylogeny of the plant receptor-like kinase (RLK) gene family and expression analysis of wheat RLK genes in response to biotic and abiotic stresses. BMC Genomics. 2023;24(1):224.
93. Rodrigues CM, de Souza AA, Takita MA, Kishi LT, Machado MA. RNA-seq analysis of *Citrus reticulata* in the early stages of *Xylella fastidiosa* infection reveals auxin-related genes as a defense response. BMC Genomics. 2013;14(1):676.
94. Bastianel M, Pereira-Martin JA, Novelli VM, Freitas-Astúa J, Nunes MA. Citrus leprosis resistance within the citrus group. Virusdisease. 2018;29(4):491–8.
95. Biswas KK, Palchoudhury S, Chakraborty P, Bhattacharyya UK, Ghosh DK, Debnath P, et al. Codon usage bias analysis of *Citrus tristeza* virus: higher codon adaptation to *Citrus reticulata* host. Viruses. 2019. https://doi.org/10.3390/v11040331.
96. Díaz L, Del Río JA, Ortuño A. Mechanism of the Alternaria alternata Pathogenicity in 'Fortune' Mandarin. Horticulturae. 2018;4(4):54.
97. Zhou CLE, Ammar E-D, Sheta H, Kelley S, Polek M, Ullman DE. Citrus tristeza virus ultrastructure and associated cytopathology in *Citrus sinensis* and *Citrus aurantifolia*. Can J Bot. 2002;80(5):512–25.
98. Barkley NA, Roose ML, Krueger RR, Federici CT. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). Theor Appl Genet. 2006. https://doi.org/10.1007/s00122-006-0255-9.
99. Herrero R, Asíns MJ, Carbonell EA, Navarro L. Genetic diversity in the orange subfamily Aurantioideae. I. Intraspecies and intragenus genetic variability. Theor Appl Genet. 1996;92(5):599–609.
100. Yu Y, Chen C, Huang M, Yu Q, Du D, Mattia MR, et al. Genetic Diversity and Population Structure Analysis of Citrus Germplasm with Single Nucleotide Polymorphism Markers. J Amer Soc Hort Sci. 2018;143(6):399–408.
101. Bhandari HR, Bhanu AN, Srivastava K, et al. e: Assessment of genetic diversity in crop plants - an overview. Adv Plants Agric Res. 2017;7(3):279–86.
102. Nelson R, Wiesner-Hanks T, Wisser R, Balint-Kurti P. Navigating complexity to breed disease-resistant crops. Nat Rev Genet. 2018;19(1):21–33.
103. Salgotra RK, Chauhan BS. Genetic diversity, conservation, and utilization of plant genetic resources. Genes. 2023. https://doi.org/10.3390/genes14010174.
104. Prigozhin DM, Krasileva KV. Analysis of intraspecies diversity reveals a subset of highly variable plant immune receptors and predicts their binding sites. Plant Cell. 2021;33(4):998–1015.
105. Wang D, Sha Y, Hu J, Yang T, Piao X, Zhang X. Genetic signatures of plant resistance genes with known function within and between species. Genetica. 2018;146(6):517–28.
106. González VM, Aventín N, Centeno E, Puigdomènech P. Interspecific and intraspecific gene variability in a 1-Mb region containing the highest density of NBS-LRR genes found in the melon genome. BMC Genomics. 2014;15(1):1131.
107. Yang S, Jiang K, Araki H, Ding J, Yang YH, Tian D. A molecular isolation mechanism associated with high intra-specific diversity in rice. Gene. 2007;394(1–2):87–95.
108. Yang S, Gu T, Pan C, Feng Z, Ding J, Hang Y, et al. Genetic variation of NBS-LRR class resistance genes in rice lines. Theor Appl Genet. 2008;116(2):165–77.
109. Sela H, Cheng J, Jun Y, Nevo E, Fahima T. Divergent diversity patterns of NBS and LRR domains of resistance gene analogs in wild emmer wheat populations. Genome. 2009;52(6):557–65.

Publisher's Note