https://doi.org/10.17221/131/2021-PPS

Genome-wide characterisation, evolution and expression analysis of the leucine-rich repeat receptor-like kinase (*LRR-RLK*) gene family in cucumbers

*Jia Yu, Bo Zhang, Sisi Liu, Wei Guo, Yifan Gao, Hongyan Sun**

College of Chemical and Biological Engineering, Taiyuan University of Science and Technology, Taiyuan, P.R. China

**Corresponding author: hongyansun@tyust.edu.cn*

Citation: Yu J., Zhang B., Liu S.S., Guo W., Gao Y.F., Sun H.Y. (2022): Genome-wide characterisation, evolution and expression analysis of the leucine-rich repeat receptor-like kinase (*LRR-RLK*) gene family in cucumbers. Plant Protect. Sci., 58: 125–138.

Abstract: The leucine-rich repeat receptor-like kinases (*LRR-RLKs*) compose a large gene family in plant genomes and implement essential functions in diverse plant physiology progress, including defence against pathogens. However, a systematic analysis of *LRR-RLKs* has not been accomplished in the economically important cucumber. 189 *LRR-RLK* genes were identified in the cucumber genome and further divided into 22 subgroups based on the sequence similarities in this study. A total of 31 segmental duplication events and 15 tandem duplication events were present in the genome, indicating that the two duplications were the main driving forces for the expansion of the *LRR-RLK* family in the cucumber. The expression profile analysis revealed that most of the *CsLRR-RLKs* were upregulated during a downy mildew infection, and resistant cucumbers comprised more upregulated *CsLRR-RLKs* than the sensitive lines. Taken together, our results provided information on the *LRR-RLK* gene family in the cucumber and contributed valuable information for the further research of *CsLRR-RLKs*.

Keywords: *Cucumis sativus*; evolutionary analysis; gene expression; genome sequence; *Pseudoperonospora cubensis*

In the PAMP (pathogen-associated molecular pattern)-triggered immunity (PTI) of plants, the microbe-associated molecular patterns (MAMPs) released from diverse pathogens as well as damage-associated molecular patterns (DAMPs) are recognised by plasma membrane-localised pattern recognition receptors (PRRs) (Macho & Zipfel 2014). PRRs include receptor-like proteins (RLPs) and receptor-like protein kinases (RLKs). RLKs are composed of an ectodomain (ECD) and an intracellular kinase domain (KD), connected by a single-pass transmembrane domain (TM) (Shiu & Bleecker 2001). RLKs are categorised into different subfamilies according to the ECD structure, and the LRR-RLK family occupies the largest number in several RLK superfamilies (Shiu & Bleecker 2001).

LRR-RLKs play an essential role in a wide array in a plant's physiology progress including microbe sensing, hormonal signals, and immune responses (Tanaka et al. 2014). *Arabidopsis* FLS2 and the EF-Tu receptor (EFR) belong to the LRR-RLK family which respectively plays a role in the recognition of the conserved bacterial flagella protein flg22 and the protein EF-Tu (Chinchilla et al. 2006;

Supported by the Key Research and Development Project of Shanxi Province (Project No. 201903D221066), Scientific and Technological Innovation Projects of Colleges and Universities in Shanxi Province (Project No. 2020L0369), Reward Fund for Outstanding Doctors Working in Shanxi (Project No. 20212017), Natural Science Foundation of Shanxi Province (Project No. 20210302124513), and the Doctoral Initiation Fund of Taiyuan University of Science and Technology (Project No. 20192057).

https://doi.org/10.17221/131/2021-PPS

Zipfel et al. 2006). The interaction between FLS2 and flg22, as well as ERF and EF-Tu, trigger the plant innate immunity including the activation of mitogen-activated protein kinase (MAPK) cascades and the expression of pathogen-related (*PR*) genes (Zipfel 2008; Yu et al. 2017). Similar to FLS2 and EFR, Xanthomonas resistance21 (Xa21) responds to sulfated peptide RaxX in rice (Luu et al. 2019). New LRR-RLKs were identified, such as xanthine/uracil permease sensing 1 (XPS1) (Mott et al. 2016). In addition to the perception of MAMPs, plant LRR-RLKs participate in recognising DAMPs as well. PEP receptor1 (PEPR1) and PEPR2 recognise endogenous PEPs (plant elicitor peptides) in *Arabidopsis* cells (Liu et al. 2013). In addition, LRR-RLKs are involved in the symbiotic relationship between rhizobia and legumes. Legumes regulate the number of nodules through a root-shoot signalling pathway called autoregulation of nodulation (AON). CLAVATA3/ESR-related (CLE) peptides are produced in the roots and are delivered to the shoot where they are recognised by LRR-RLKs such as HYPERNODULATION ABERRANT ROOT FORMATION 1 (HAR1), KLAVIER (KLV) and SUPER NUMERIC NODULES (SUNN) (Miyazawa et al. 2010; Tiwari et al. 2021). Another LRR-RLK, Compact Root Architecture 2 (CRA2), is identified to systemically regulate the nodule initiation (Huault et al. 2014; Zhu et al. 2020). Furthermore, studies have revealed the role of LRR-RLKs during abiotic stress including drought and salinity stress (Wang et al. 2017; Yuan et al. 2018). PnLRR-RLK27, an LRR-RLK gene in *Pohlia nutans*, is supposed to positively regulate the oxidative-stress and salinity tolerance (Wang et al. 2017). A recent transcriptome analysis indicated that several LRR-RLKs have a potential function during water deficiency in *Thinopyrum elongatum* (Mishra et al. 2021).

Given the crucial functions of *LRR-RLKs* in plants, this gene family has been discussed in diverse kinds of plants in addition to the typical model plant *Arabidopsis thaliana* (Hwang et al. 2011). For instance, *LRR-RLKs* in some important species of Solanaceae (Li et al. 2018), Brassicaceae (Wang et al. 2019), Rosaceae (Sun et al. 2017) and Leguminosae (Zhou et al. 2016; Meng et al. 2020) have been estimated by a genome-wide analysis along with some woody species, such as the Populus (Zan et al. 2013) and citrus (Magalhães et al. 2016). Recent research studies have paid attention to *LRR-RLKs* related to biotic stress responses by using genome and transcriptome analyses (He et al. 2019; Cao et al. 2021). Consequently, the utilisation of bioinformatic methods to systematically investigate the evolution and expression profiles of the *LRR-RLKs* contributes to gain knowledge of the possible functions of the *LRR-RLK* genes in regulating plant innate immunity.

The cucumber (*Cucumis sativus*), which belongs to Cucurbitaceae, is an economically important vegetable crop and serves as a model system to research plant defence responses to pathogens. Downy mildew caused by the biotrophic pathogen *Pseudoperonospora cubensis* is one of the important limiting factors in cucumber cultivation and production. *P. cubensis*, an oomycete pathogen, is capable of infecting other genera within Cucurbitaceae causing significant losses. Potential pathways in the defence against downy mildew infections of the cucumber were revealed in previous studies including the salicylic acid-mediated signalling pathway (Yan et al. 2020). However, the molecular regulation mechanisms of the cucumber resistance to downy mildew are poorly known. The evolutionary relationships, possible gene functions and mechanisms of the *LRR-RLK* gene family during defence against downy mildew remain to be further understood. Based on public genome and transcriptome sequencing data, we carried out a systematic analysis of the *LRR-RLK* genes of the cucumber in this study. The study aimed to clarify the phylogenetic relationships of *LRR-RLKs* as well as the duplication events and expression patterns. Furthermore, the determination of the transcriptional activity of these genes during a pathogen infection will provide insight into the *LRR-RLKs*' function.

MATERIAL AND METHODS

Identification of *LRR-RLKs* in *C. sativus* genome. The genome and annotation resources of *A. thaliana* were acquired from the TAIR10 database (http://www.arabidopsis.org). The cucumber genome (Chinese Long v3) and annotations were acquired from the GuGen database (http://cucurbitgenomics.org). The Hidden Markov Model (HMM) profiles of the LRR domains including PF00560, PF07723, PF07725, PF12799, PF13306, PF13516, PF13855, PF14580 and PF01816 were downloaded from the PFAM database (https://pfam.xfam.org),

https://doi.org/10.17221/131/2021-PPS

followed by an HMM search (*E*-value cut-off < 1) against the cucumber genome to identify the LRR domain containing proteins. The HMM profile of the kinase domains (PF00069 and PF07714) were also scanned (*E*-value cut-off < $1e^{-5}$). In addition, the LRR-RLKs identified in *Arabidopsis*, rice, soybean and Poplar (Sun & Wang 2011; Zan et al. 2013; Zhou et al. 2016; Meng et al. 2020) were applied to perform the BlastP against the *C. sativus* proteome database (*E*-value < 10^{-5}) (Camacho et al. 2009). All of the obtained sequences were further checked using the conserved domain database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and InterPro (https://www.ebi.ac.uk/interpro). The Transmembrane Helices; Hidden Markov Model (TMHMM) website (http://www.cbs.dtu.dk/services/TMHMM) was applied to predict the transmembrane domain.

Phylogenetic analysis and chromosomal location of *CsLRR-RLKs*. To access the evolutionary and classification of the *C. sativus LRR-RLK* genes, a phylogenetic tree was established via MEGA (version 7.0) using a neighbour-joining (NJ) algorithm with 1 000 bootstrap replicates (Kumar et al. 2016). The phylogenetic tree was built with the aligned KD sequences of 189 *C. sativus* and 225 *A. thaliana LRR-RLKs* (Meng et al. 2020). Clustal X (version 2.0.11) was applied in multiple sequence alignments with default parameters (Larkin et al. 2007). Physical maps of the *CsLRR-RLK* genes were constructed showing the genome localisation and chromosome information collected from the genomic annotation files. The visualisation of the *LRR-RLKs* on seven chromosomes was realised by using MG2C (http://mg2c.iask.in/mg2c_v2.0).

Structural organisation of *LRR-RLKs* and conserved motifs analysis. The genome sequence and cDNA sequences were downloaded from the GuGen database (Burkhardt & Day 2016) which corresponded to each predicted *CsLRR-RLKs* gene annotation database. Then the exons distribution pattern was examined using the online analysis tool GSDS (version 2.0) (http://gsds.gao-lab.org). To further evaluate the evolution of the *LRR-RLK* genes in *C. sativus*, the conserved motifs of the KDs were identified using the Multiple EM for Motif Elicitation (MEME) online program (version 5.3.3) (https://meme-suite.org/meme/tools/meme) (Bailey et al. 2006).

Gene duplication and synteny analysis. To accesses the generation of *CsLRR-RLKs*, a tandem duplication and synteny analysis of the *C. sativus* and *A. thaliana LRR-RLK* genes were conducted. The tandem duplications of *CsLRR-RLKs* were characterised as genes closely clustered within 200 kb. The synteny analysis of the *LRR-RLKs* from *C. sativus* and *A. thaliana* was mapped by MCScanX (Wang et al. 2012) in addition to the identification of the tandem and segmental duplications.

Expression analysis of the *C. sativus LRR-RLK* genes under infection. To estimate the expression patterns of the *CsLRR-RLKs* during a pathogen infection, RNA-seq data (PRJNA285071) were collected from the GuGen database (Burkhardt & Day 2016). Susceptible (Vlaspik) and resistant (PI 197088) cucumber lines were inoculated with downy mildew, *Pseudoperonospora cubensis*. Twelve different data sets were selected that contained data from two uninfected cucumber lines and in response to a time course (1, 2, 3, 4, and 6 dpi) of infection with *P. cubensis*. The RNA-sequencing reads were aligned and mapped to the cucumber genome utilising the HISAT2 software (Expósito et al. 2018). Then expression level was evaluated from the fragments per kilobase of transcript per million mapped reads values using StringTie with default parameters (Pertea et al. 2015). A differentially expressed gene (DEG) analysis was carried out using the edgeR program (Robinson et al. 2010).

RESULTS

Genome-wide identification of *LRR-RLK* genes of *C. sativus*. In this study, only proteins that comprised at least one LRR domain, one transmembrane helix, and one kinase domain were treated as the LRR-RLK protein. Based on the HMM search approach, 189 *C. sativus* genes were classified into the *LRR-RLK* gene family which accounted for 0.78% of the total protein-coding genes [Figure 1; Table 1; Table S1 in electronic supplementary material (ESM); for the supplementary material see the electronic version].

The proportion of *LRR-RLKs* was largely in accord with previous results which indicated 0.67–1.39% proportions in angiosperm species (Liu et al. 2017). To explore the genomic distributions of the *C. sativus LRR-RLK* genes in the chromosomes, the genomic position of each *LRR-RLK* gene was identified. A total of 179 genes were mapped to all seven chromosomes of *C. sativus*, and the oth-

https://doi.org/10.17221/131/2021-PPS

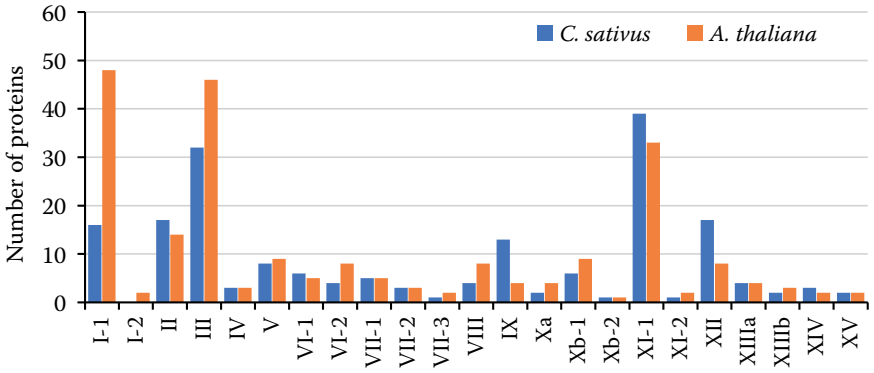

Figure 1. The distribution of the *LRR-RLK* gene family members in 23 subgroups in *Cucumis sativus* and *Arabidopsis thaliana*

er ten members were allotted to unassembled scaffolds of the *C. sativus* genome which could not be localised to any chromosome (Figure 2). The overall distribution of the *CsLRR-RLKs* on the chromosomes tended to be unequal. The distribution proportion on the seven chromosomes indicated the gene localisation varied from 8.47% (16 members on chromosome 2) to 21.69% (41 genes on chromosome 3) (Table S2 in ESM). More than half (51.85%) of the *LRR-RLK* genes were located on chromosomes 1, 3 and 5.

Phylogenetic analysis and classification of the *CsLRR-RLK* genes. To evaluate the evolutionary characteristics of the *C. sativus LRR-RLK* genes, the kinase domains (KDs) were firstly retrieved from each LRR-RLK protein. The evolutionary tree

Table 1. Total number of genes and exons distributed in the *LRR-RLK* different subgroups

Subgroups	*Arabidopsis thaliana*	*Cucumis sativus*	Maximum number of exons	Minimum number of exons	Average number of exons
I-1	48	16	16	7	12
I-2	2	0	–	–	–
II	14	17	12	10	10.94
III	46	32	4	2	2.31
IV	3	3	4	4	4
V	9	8	16	14	15.63
VI-1	5	6	7	2	6.17
VI-2	8	4	13	2	9
VII-1	5	5	3	2	2.2
VII-2	3	3	3	1	2
VII-3	2	1	1	1	1
VIII	8	4	20	19	19.25
IX	4	13	24	2	16.77
Xa	4	2	2	1	1.5
Xb-1	9	6	2	1	1.17
Xb-2	1	1	2	2	2
XI-1	33	39	5	1	2.26
XI-2	2	1	1	1	1
XII	8	17	11	1	2.76
XIIIa	4	4	13	11	12.5
XIIIb	3	2	27	27	27
XIV	2	3	3	3	3
XV	2	2	1	1	1
Total	225	189	–	–	–

https://doi.org/10.17221/131/2021-PPS

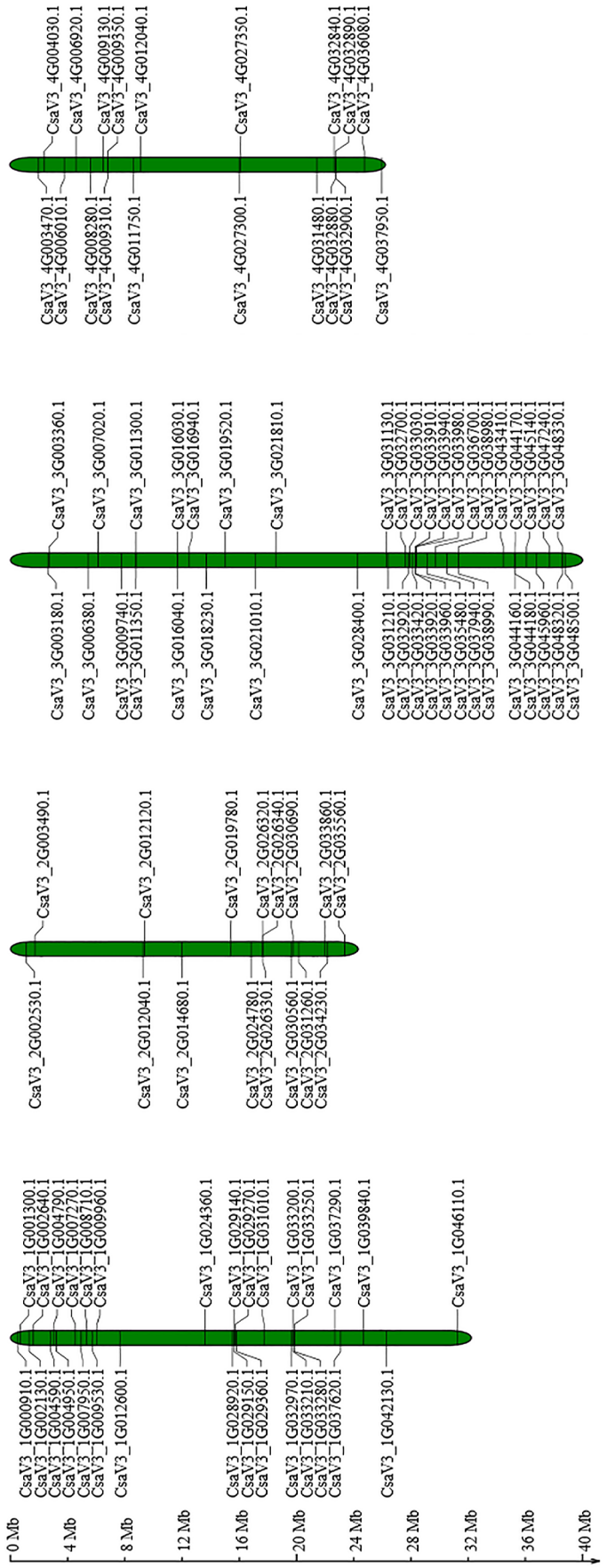

Figure 2. Distribution of the *LRR-RLK* genes on seven chromosomes of *Cucumis sativus*
The chromosomal positions of *LRR-RLK* genes were revealed according to the physical positions. Green bars indicated the chromosomes on which the black lines showed the positions of the *LRR-RLKs*

https://doi.org/10.17221/131/2021-PPS

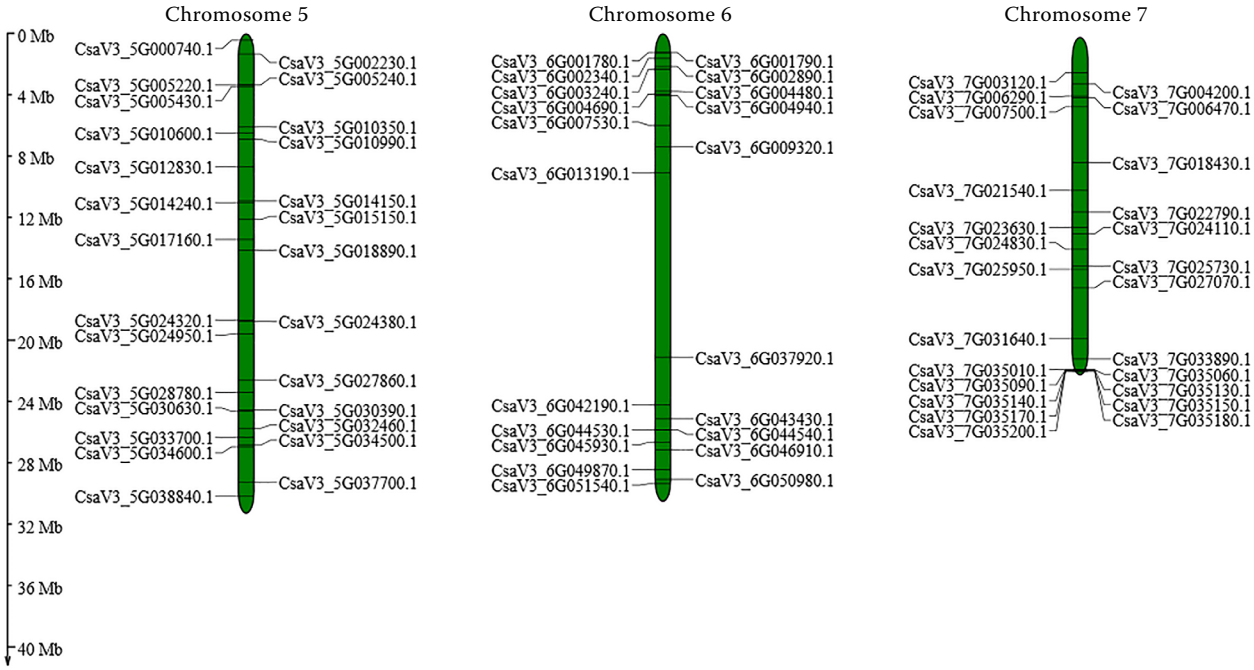

Figure 2 to be continued

was established based on all the identified *C. sativus* and *A. thaliana* KDs by the neighbour-joining algorithm (Figure 3). The sequences of 225 *A. thaliana* LRR-RLK proteins were obtained from a previous research study (Meng et al. 2020). Per the phylogenetic tree shown in Figure 3, the LRR-RLK proteins in *C. sativus* were divided into 15 groups and 22 subgroups, all of which were also present in *A. thaliana*. Notably, group I-2, identified in *A. thaliana*, was not found in *C. sativus*.

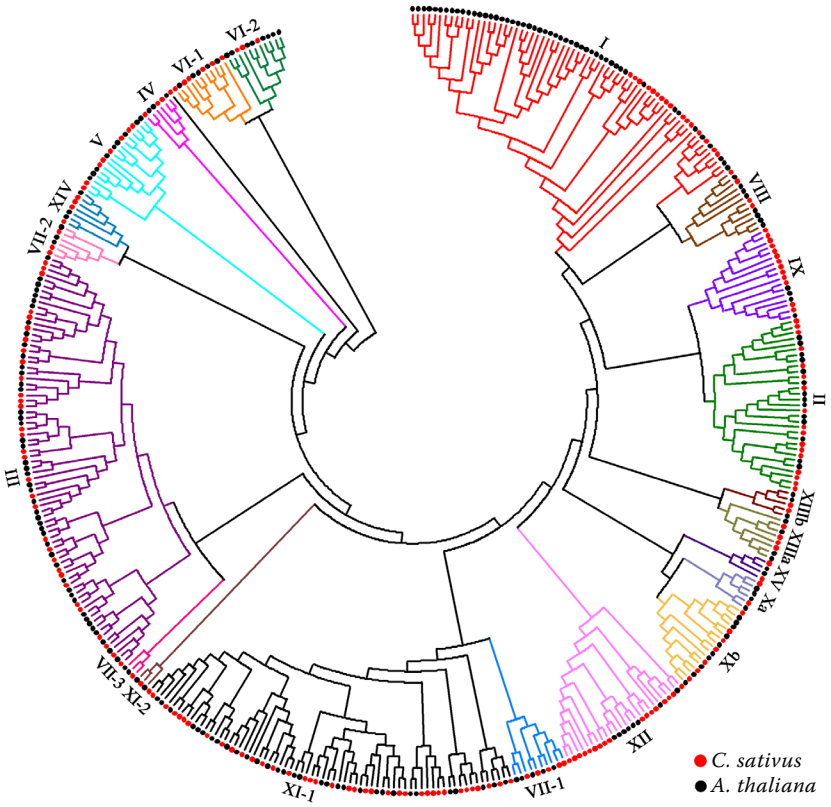

Figure 3. Phylogenetic tree for the LRR-RLKs

The phylogenetic tree was constructed based on the KDs from *Cucumis sativus* and *Arabidopsis thaliana* LRR-RLKs by the NJ method. *C. sativus* and *A. thaliana* LRR-RLK are differentiated by red and black dots. Different LRR-RLK subgroups were presented in different line colours. The reported *A. thaliana* LRR-RLKs with known functions (Wu et al. 2016; Li et al. 2017; Lu et al. 2020) were marked with a triangle of which the colour represented a different function in *A. thaliana*. The function of the plant growth, seed development, flower development, phytohormone signalling, biotic stress and abiotic stress were individually marked with green, yellow, red, blue, purple and black triangles, respectively

https://doi.org/10.17221/131/2021-PPS

The cucumber *LRR-RLK* genes were categorised into 15 groups and 22 subgroups. Each subgroup contained a different number of genes which ranged from one to 39 (Figure 3; Table 1). Notably, LRR-I contained only one subgroup in *C. sativus* while it contained two subgroups in *A. thaliana.* Among the 22 subgroups in *C. sativus*, VII-3, Xb-2, XI-2 were the smallest subgroups comprised of one gene only, and XI-1 and III were the two largest subgroups with 39 and 32 genes, individually. The other subgroups were comprised of no more than 30 members. Regarding the LRR-I group, only 16 members were identified in *C. sativus*, while 50 members were reported for *A. thaliana*, which is more than three times when compared with *C. sativus.* Group XII and IX in *C. sativus* contained more than twice the number of members than in *A. thaliana*.

Analysis of the gene structure and conserved motif in the kinase domains. To investigate the gene structures of the *LRR-RLK* genes, the intron-exon structure of each *LRR-RLK* gene was identified, followed by a further analysis (Table 1). The number of exons was varied from one to 24 among all the *CsLRR-RLKs*. Thirteen *LRR-RLK* members contained only one exon that accounted for 6.88% of the total members (Table S1 in ESM). There were 80 members with more than three exons and 10 with more than 20 exons (Table S1 in ESM). In addition to the phylogenetic analysis, it was suggested that members in the same subgroups tended to contain similar numbers of exons. For example, the LRR-III subgroup comprised 32 members with an average of 2.31 exons and a maximum of four (Table 1). However, the exon numbers of the IX subgroup genes ranged from two to 24 with an average of 16.77, though it comprised only 17 members.

LRR-RLKs are composed of LRR, TM and kinase domains. The kinase domain of the CsLRR-RLKs consisted of an average of 246 amino acids with a maximum of 293 and a minimum of 84 amino acids. To investigate the evolutionary divergence of the kinase domain, a motif analysis was performed with the MEME program. A total of 12 motifs were recognised in the CsLRR-RLK kinase domains, and they were named M1 to M12 according to the frequency of their appearance (Figure 4). The majority of the motifs did not seem to be subfamily-specific. In comparison with other plants, the majority of the 12 motifs

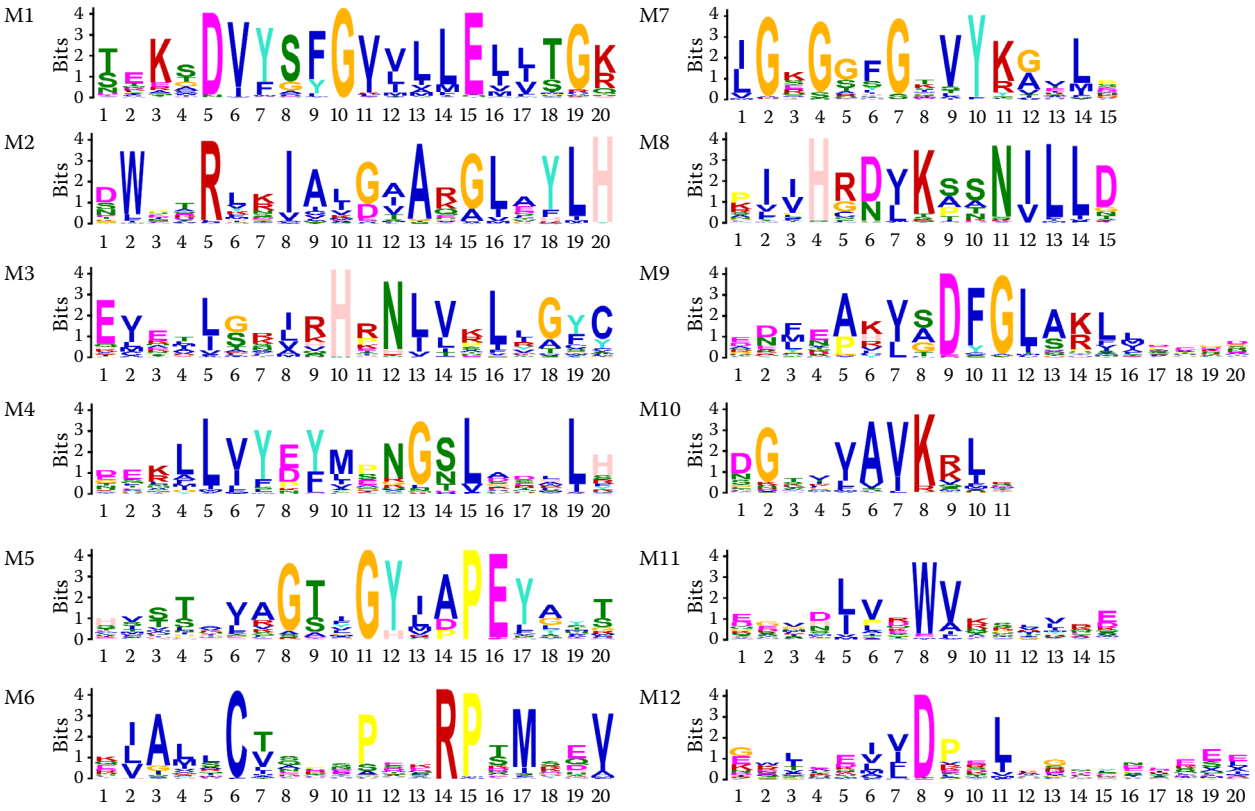

Figure 4. Conserved motifs (M1 to M12) and consensus sequences in the kinase domain of the *Cucumis sativus* LRR-RLK proteins

https://doi.org/10.17221/131/2021-PPS

were highly similar to those identified in *Amborella trichopoda* (Liu et al. 2016), *Populus trichocarpa* (Zan et al. 2013) and *Oryza sativa* (Sun & Wang 2011), suggesting that the kinase domains of the LRR-RLKs may play conserved functions among different plants.

Gene duplication and synteny of the *LRR-RLK* genes in *C. sativus*. Gene duplication is referred to as a kind of main driving power for promoting genome evolutions. In this study, tandem duplication events referred to a region on a chromosome which comprised two or more genes within 200 kb. Fifteen transcriptional diminished (TD) gene pairs involving 41 genes were distributed across all seven chromosomes (Table S3 in ESM). There were as many as six TD gene pairs involving 16 genes on chromosome 3 which comprised 39.02% of all the TD gene pair members (Table S3 in ESM). Among the 15 TD gene pairs, the largest two clusters had six and five tightly linked genes which individually occurred on chromosomes 7 and 3 (Table S3 in ESM), respectively. Most TD events were distributed in the LRR-XII and LRR-XI-1 subgroups, each of which contained five clusters with 14 and 10 genes, separately. Importantly, the XI-1 subgroup had the most members in *C. sativus*. The number of the XII genes in *C. sativus* was twice as many as that in *A. thaliana* (Table 1). These results indicated that the TD events in *C. sativus* are one reason to clarify the increase of members of these two subgroups.

The segmental duplication events occurring in the *CsLRR-RLK* gene family were evaluated by conducting a synteny analysis of the genes using MCScanX. As shown in Figure 5, the results indicated that the *C. sativus* genome contains many segmental replication events. A total of 31 pairs of segmental duplications in the *CsLRR-RLK* gene family were identified to be collinearly related (Table S3 in ESM). Notably, the majority of the *CsLRR-RLK* genes underwent segmental duplication were located in chromosome 1, 3 and 5. Additionally, the synteny of the *LRR-RLKs* between the *C. sativus* and *A. thaliana* genomes were analysed to further estimate the phylogenetic mechanisms of the *CsLRR-RLKs* (Figure 6). A total of 175 pairs of *LRR-RLK* genes were present in the syntenic blocks between *C. sativus* and *A. thaliana*, which highlighted the common ancestral origin (Table S3 in ESM). Moreover, 49 *CsLRR-RLKs* were not included in any syntenic blocks, suggesting that these *LRR-RLKs* may originate from independent duplication events.

Expression profiles of the *CsLRR-RLKs* in response to the downy mildew infection. To acquire a broader knowledge of the functions of *CsLRR-RLKs*, an expression profile analysis was applied based on the public RNA-seq data. The expression patterns of all 189 *CsLRR-RLKs* were assessed based on the transcriptomic data of susceptible (Vlaspik) and resistant (PI 197088) cucumber lines in response to a time course (1, 2, 3, 4, and 6 dpi) of infection with *P. cubensis*. A total of 171 *CsLRR-RLKs* were expressed at diverse infection time points of both two plant lines according to the transcriptome data (Figure 7; Table S4 in ESM).

The *LRR-RLK* genes presented different expression profiles in response to the downy mildew infection as shown in Figure 7 and Table S4 in ESM. It was suggested that most of the *CsLRR-RLKs* were up-regulated in response to the infection both in the susceptible and resistant cucumber lines. Among the up-regulated *CsLRR-RLKs*, some sustained high expression levels in two lines, for instance CsaV3_3G044160.1. The up-regulated fold of CsaV3_3G044160.1 in the susceptible line was much higher than in the resistant line. Moreover, the highest expression level of CsaV3_7G035010.1 came earlier in the resistant line than in the susceptible line. CsaV3_3G044160.1 and CsaV3_7G035010.1 both belong to subgroup I. Overall, the number of up-regulated *LRR-RLK* genes in PI 197088 (resistant) was more than in Vlaspik (susceptible).

A total of 105 DEGs were distributed in 19 subgroups except for VII-3, XII and XV. Subgroup XI-1, III and I individually comprised 20, 18 and 12 differentially expressed *CsLRR-RLKs* which contained nearly 50% of all the DEGs in total. Although the number of *CsLRR-RLKs* in I were reduced to 1/3 compared with *A. thaliana*, 75% of the genes of this subgroup inhibited different expression levels. In addition, most *LRR-RLK*-I DEGs were significantly up-regulated during the downy mildew infection. However, the DEGs in II were mostly down-regulated with significance in response to *P. cubensis*. The diverse expression profiles of *CsLRR-RLKs* during the infection suggested the complicated roles of *CsLRR-RLKs* in response to *P. cubensis*.

DISCUSSION

The *LRR-RLK* gene family that is involved in the stress responses has been characterised in many

https://doi.org/10.17221/131/2021-PPS

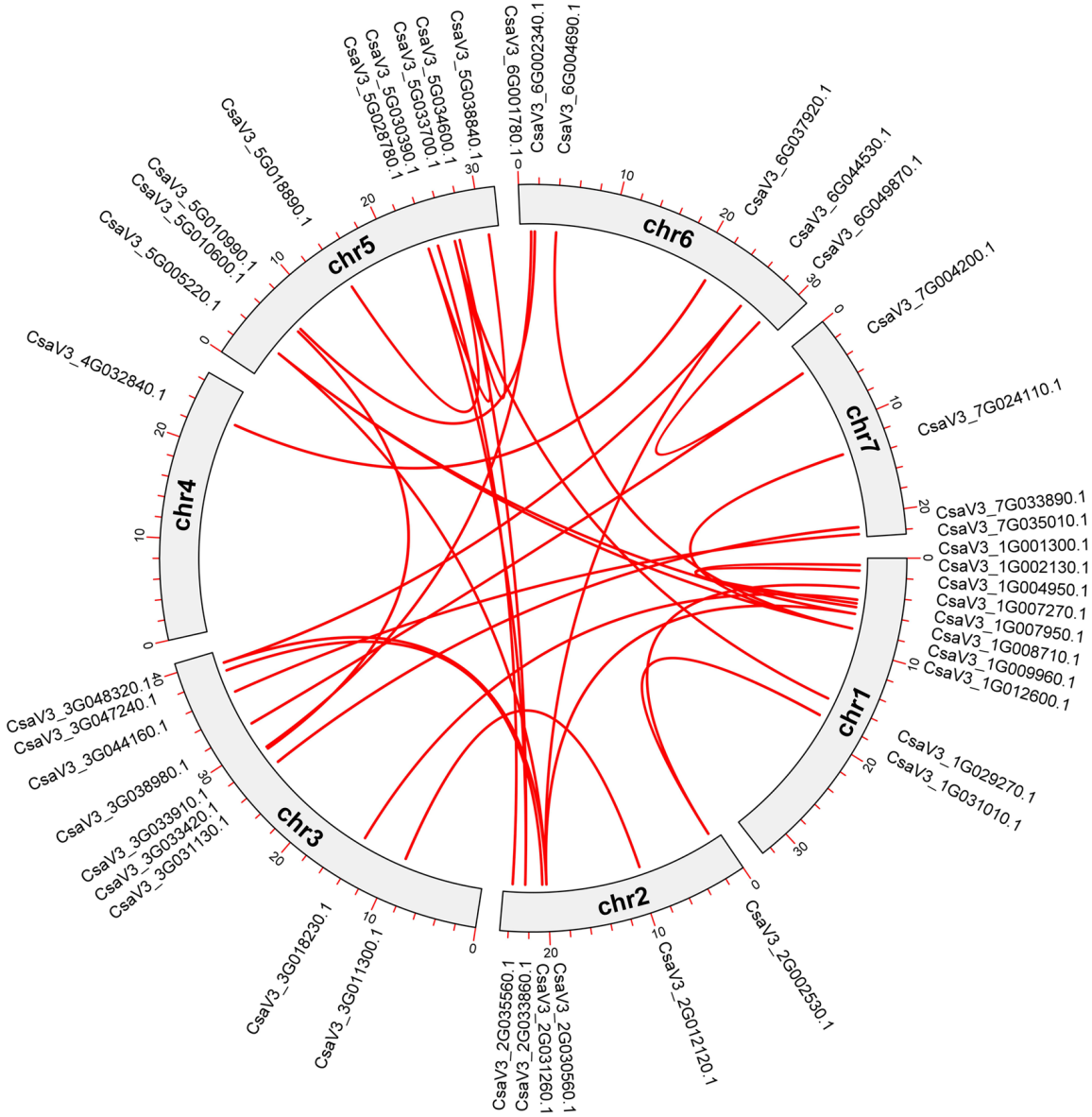

Figure 5. Chromosome localisation of the segmental duplication gene pairs of *LRR-RLKs*
The red lines revealed the segmented duplicated gene pairs

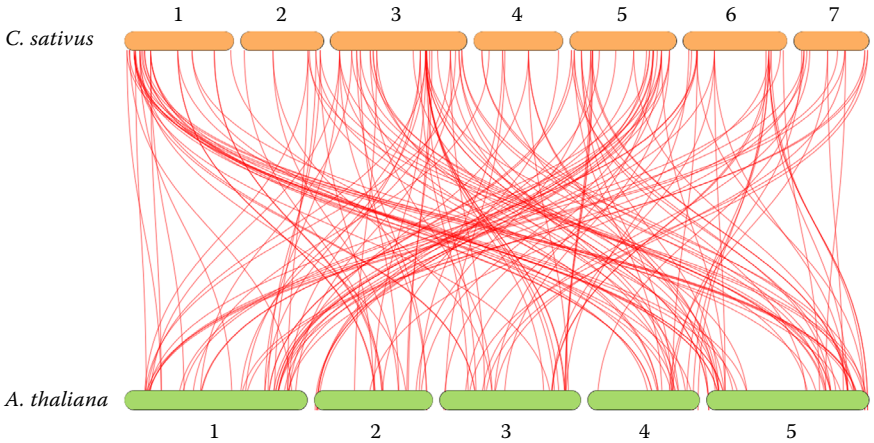

Figure 6. The synteny of the *LRR-RLK* genes in the genomes of *Cucumis sativus* and *Arabidopsis thaliana*

https://doi.org/10.17221/131/2021-PPS

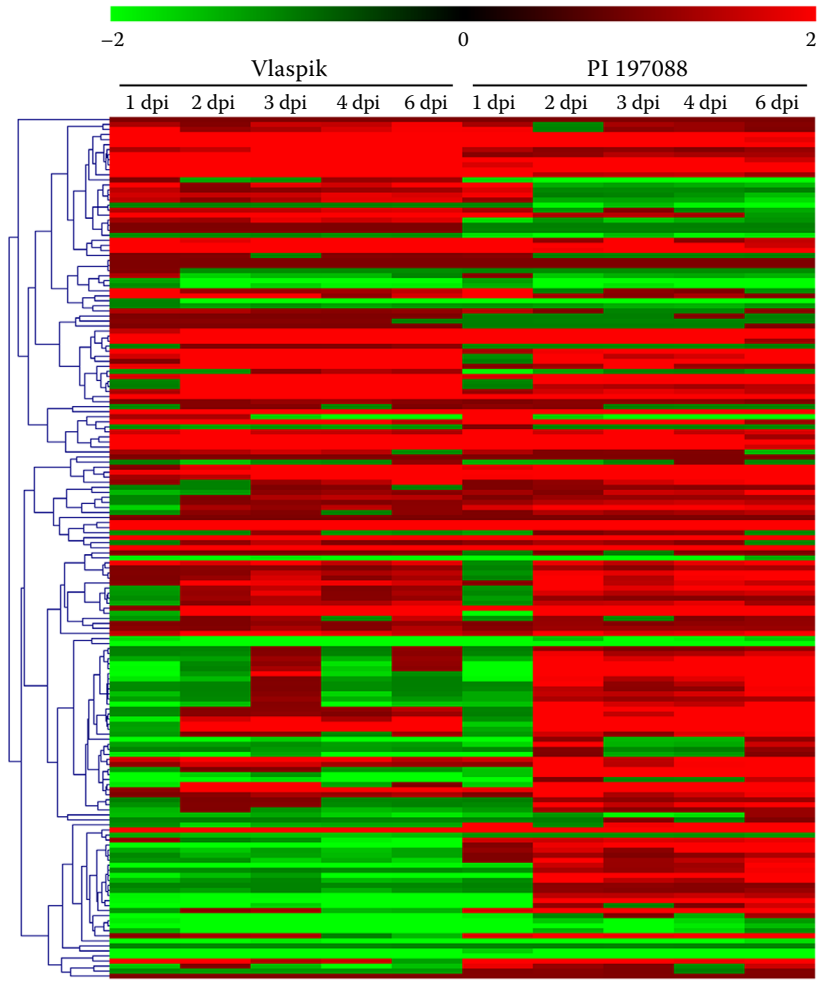

Figure 7. Expression pattern of the *CsLRR-RLK* genes in responding to the downy mildew infection
dpi – days post inoculation
The heatmap depicts the expression profiles of the *CsLRR-RLKs* in the susceptible (Vlaspik) and resistant (PI 197088) lines of cucumbers in response to a time course of infection with downy mildew

plant species (Li et al. 2018; Meng et al. 2020). However, the characteristics and functions of the *C. sativus LRR-RLK* genes remain unclear. In this study, the whole-genome scanning revealed that there were 189 genes considered to be putative *LRR-RLK* genes. The number of *LRR-RLK* family members in *C. sativus* occupied 0.78% of the whole-genome coding proteins, which accorded to the proportions of 0.67–1.39% in angiosperm species (Liu et al. 2017). The *LRR-RLK* gene family is a comparatively large family which has diversified to hundreds of members in many plants. In comparison with other dicot genomes (Table 2), the proportion of the *LRR-RLKs* in the genomic protein-coding genes of *C. sativus* was similar to *B. rapa* (0.75%) and *A. thaliana* (0.82%). The number of protein-coding genes in the whole genome was similar in *C. sativus*, *C. clementina* and *C. sinensis*, while the proportion of the *LRR-RLKs* was significantly distinct (Table 2). These *CsLRR-RLK* genes were further classified into 22 subgroups by comparing the homology of the *Arabidopsis LRR-RLKs*. The proportion of subgroups were varied from diverse plants. The number of LRR-I subgroup members occupied 8.47% in *C. sativus* which was equal to *P. trichocarpa* (8.71%) (Table 2). The percentage of cucumber LRR-XII members (8.99%) was similar to *B. rapa* (8.20%). While the LRR-IX members in *C. sativus* (6.88%) was nearly four times the proportion of *A. thaliana* (1.78%).

It was demonstrated that *LRR-RLKs* might have a particular tendency towards genomic recombination events for repetitive LRR domains, a high copy number as well as the high ratio of defence-related genes (Hofberger et al. 2014). Tandem and segmental duplication were regarded as the important reasons for the expansion of gene families for the vast majority of plant species. A total of 15 TD gene pairs and 31 pairs of segmental duplications were identified in the *C. sativus* genome which were distributed across all seven chromosomes (Figure 5; Table S3 in ESM). The distribution of the duplicated *LRR-*

https://doi.org/10.17221/131/2021-PPS

Table 2. Total number of *LRR-RLKs*, LRR-I and LRR-XII in the plants

Plant species	Number of protein-coding genes	*LRR-RLKs*	LRR-I	LRR-IX	LRR-XII	References
Vernicia fordii	28 422	167 (0.59%)	7 (4.19%)	7 (4.19%)	41 (24.55%)	Cao et al. 2021
Glycine max	56 044	467 (0.83%)	23 (4.93%)	16 (3.43%)	73 (15.63%)	Zhou et al. 2016
Medicago truncatula	44 623	329 (0.74%)	23 (6.99%)	11 (3.34%)	65 (19.76%)	Meng et al. 2020
Populus trichocarpa	41 335	379 (0.92%)	33 (8.71%)	12 (3.17%)	42 (11.08%)	Zan et al. 2013
Citrus clementina	24 533	300 (1.22%)	9 (3.00%)	6 (2.00%)	148 (2.00%)	Magalhães et al. 2016
Citrus sinensis	25 376	297 (1.17%)	11 (3.70%)	5 (1.68%)	140 (47.14%)	Magalhães et al. 2016
Brassica rapa	40 492	305 (0.75%)	41 (13.44%)	11 (3.61%)	25 (8.2%)	Wang et al. 2019
Raphanus sativus	53 642	292 (0.54%)	46 (15.75%)	8 (2.74%)	23 (7.88%)	Wang et al. 2019
Arabidopsis thaliana	27 416	225 (0.82%)	48 (21.33%)	4 (1.78%)	8 (3.56%)	Meng et al. 2020
Cucumis sativus	24 317	189 (0.78%)	16 (8.47%)	13 (6.88%)	17 (8.99%)	this article

RLKs was uneven both on the chromosomes and related subgroups. It was suggested that the TD events were distributed in the subgroup XII and XI-1. LRR-XII comprised the most TD gene pairs and twice the number of members than in *A. thaliana*. The phylogenetic tree in Figure 3 indicates that most cucumber LRR-XI-1 genes had homologous genes in *Arabidopsis.* Cucumber genes in the XII subgroup was homogenous with limited *Arabidopsis* genes in this group. The result indicated that tandem duplications played an essential role in the expansion of this subgroup. It was suggested that the RLK genes, which participated in responding to different stresses, tended towards being duplicated through tandem duplication, which potentially resulted in the gene functional redundancy. In addition, gene duplication occurred in subgroup IX in cucumbers. Several *Arabidopsis* genes in subgroup IX were reported to function in the phytohormone signalling pathways (Figure 3). The function of this subgroup of genes in the cucumber needs to be further studied. The phylogenetic tree indicated that loss of *CsLRR-RLK* genes occurred in subgroup I. In subgroup I, the cucumber genes had high homology to the *Arabidopsis* genes that were related to the biotic stress and plant growth. *CsLRR-RLK* genes genes may have lost that homology to *Arabidopsis* with unknown functions in subgroup I.

The majority of *Arabidopsis LRR-RLKs* with comparable functions tended to cluster together (Shiu & Bleecker 2001), which indicated that the homologous genes of the cucumber in the same subgroup with *A. thaliana* might function similarly. For instance, group XI included members related to regulate the calcium signals, stomatal movement and plant immunity (Yang et al. 2017). The *Arabidopsis* SRF3 that belonged to group V was proven to take part in plant immune responses to pathogens (Alcázar et al. 2010). A potato LRR-RLK protein positively regulated the plant immunity to an oomycete infection that showed high sequence similarity to *Arabidopsis* SRF3 (Wang et al. 2018). Additionally, group X members played vital roles in the brassinosteroid signal transduction which participated in the regulation of the growth, development as well as innate immunity in plants (Lozano-Elena & Caño-Delgado 2019). As shown in Figure 3, previous studies revealed the functions of *Arabidopsis* LRR-RLKs in the diverse progress of plants including plant growth, seed development, flower development, phytohormone signalling, biotic stress and abiotic stress (Wu et al. 2016; Li et al. 2017; Lu et al. 2020). The study of *CsLRR-RLKs* evolutionary relationships and *Arabidopsis* homologous genes will contribute to reveal their roles in cucumbers.

Cucumber downy mildew was one of the most significant foliar diseases. The detailed insight of the molecular mechanisms associated with the plant resistance was currently restricted in the cucumber compared to *Arabidopsis* and rice. Previous studies showed that LRR-RLKs were MAMP or DAMP receptors, which play significant roles in a plant's immunity (Chinchilla et al. 2006; Wang et al. 2016). Another LRR-RLK, BAK1, one of the subgroup II members, functioned as a coreceptor in diverse PRR complexes (Sun et al. 2013). The expression profiles of *CsLRR-RLKs* in response to a downy mildew infection were analysed in this study. CsaV3_3G044160.1 and CsaV3_7G035010.1 that belonged to subgroup I were up-regulated in both the resistant and sus-

https://doi.org/10.17221/131/2021-PPS

ceptible line during the infection of the pathogen. However, the expression pattern of these two genes were distinct in the resistant and susceptible line. In the phylogenetic tree, these two LRR-RLKs were close to two *Arabidopsis* LRR-RLKs, IOS1 (Impaired Oomycete Susceptibility 1, At1g05700) and FRK1 (Flg22-induced Receptor-like Kinase 1, At2g19190) which were reported in previous research studies. IOS1 (Impaired Oomycete Susceptibility 1), a member of *Arabidopsis* LRR-I group, was required for the promotion of diseases caused by (hemi)biotrophic filamentous oomycetes along with fungal pathogens (Hok et al. 2014). Another study found that *Arabidopsis* IOS1 was important to the defence of bacteria *Pseudomonas syringae* and the priming of PTI activation (Yeh et al. 2016). FRK1 is a flg22-induced gene involved in early defence signalling (Asai et al. 2002). The plant *LRR-RLK* gene family had been extensively discussed which has a crucial role in stress responses. However, the functions of abundant *LRR-RLK* genes are still largely unknown. The functions of *C. sativus LRR-RLKs* remain to be further studied as well.

CONCLUSION

In this study, *LRR-RLK* genes were identified in the cucumber genome and were further evaluated to determine the diversity of members, evolutionary relationships, duplication events, and expression profiles during responding to the infection of downy mildew. The particular expansion was discovered in group XII compared to *Arabidopsis* which might be associated with the tandem duplication. As stated by the expression profiles analysis, the great majority of the *CsLRR-RLKs* were up-regulated during the *P. cubensis* infection both in the susceptible and the resistant lines. Furthermore, the amount of up-regulated *LRR-RLK* genes in the resistant line was more than in the susceptible line, which indicated that *LRR-RLKs* might have an important function in the response to the downy mildew infection. The results in this study provided an understanding of the evolution and potential functions of cucumber *LRR-RLK* genes and contributed to the further functional investigation.

Acknowledgement: We thank Dr. Danyu Shen (Department of Plant Pathology, Nanjing Agricultural University) for revising the manuscript.

REFERENCES

Alcázar R., García A.V., Kronholm I., de Meaux J., Koornneef M., Parker J.E., Reymond M. (2010): Natural variation at Strubbelig Receptor Kinase3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. Nature Genetics, 42: 1135–1139.

Asai T., Tena G., Plotnikova J., Willmann M.R., Chiu W.L., Gomez-Gomez L., Boller T., Ausubel F.M., Sheen J. (2002): MAP kinase signalling cascade in *Arabidopsis* innate immunity. Nature, 415: 977–983.

Bailey T.L., Williams N., Misleh C., Li W.W. (2006): MEME: Discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Research, 34: W369–W373.

Burkhardt A., Day B. (2016): Transcriptome and small rnaome dynamics during a resistant and susceptible interaction between cucumber and downy mildew. Plant Genome, 9. doi: 10.3835/plantgenome2015.08.0069

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T.L. (2009): BLAST+: Architecture and applications. BMC Bioinformatics, 10. doi: 10.1186/1471-2105-10-421

Cao Y., Mo W., Li Y., Li W., Dong X., Liu M., Jiang L., Zhang L. (2021): Deciphering the roles of leucine-rich repeat receptor-like protein kinases (LRR-RLKs) in response to *Fusarium* wilt in the *Vernicia fordii* (Tung tree). Phytochemistry, 185. doi: 10.1016/j.phytochem.2021.112686

Chinchilla D., Bauer Z., Regenass M., Boller T., Felix G. (2006): The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell, 18: 465–476.

Expósito R.R., González-Domínguez J., Touriño J. (2018): HSRA: Hadoop-based spliced read aligner for RNA sequencing data. PLoS One, 13. doi: 10.1371/journal.pone.0201483

He X., Feng T., Zhang D., Zhuo R., Liu M. (2019): Identification and comprehensive analysis of the characteristics and roles of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in *Sedum alfredii* Hance responding to cadmium stress. Ecotoxicology and Environmental Safety, 167: 95–106.

Hofberger J.A., Zhou B., Tang H., Jones J.D., Schranz M.E. (2014): A novel approach for multi-domain and multi-gene family identification provides insights into evolutionary dynamics of disease resistance genes in core eudicot plants. BMC Genomics, 15. doi: 10.1186/1471-2164-15-966

Hok S., Allasia V., Andrio E., Naessens E., Ribes E., Panabières F., Attard A., Ris N., Clément M., Barlet X., Marco Y., Grill E., Eichmann R., Weis C., Hückelhoven R., Ammon A., Ludwig-Müller J., Voll L.M., Keller H. (2014): The receptor kinase IMPAIRED OOMYCETE SUSCEPTIBILITY1

https://doi.org/10.17221/131/2021-PPS

attenuates abscisic acid responses in Arabidopsis. Plant Physiology, 166: 1506–1518.

Huault E., Laffont C., Wen J., Mysore K.S., Ratet P., Duc G., Frugier F. (2014): Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase. PLoS Genetics, 10. doi: 10.1371/journal.pgen.1004891

Hwang S.G., Kim D.S., Jang C.S. (2011): Comparative analysis of evolutionary dynamics of genes encoding leucine-rich repeat receptor-like kinase between rice and *Arabidopsis*. Genetica, 139: 1023–1032.

Kumar S., Stecher G., Tamura K. (2016): MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, 33: 1870–1874.

Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. (2007): Clustal W and Clustal X version 2.0. Bioinformatics, 23: 2947–2948.

Li Z., Wang Y., Huang J., Ahsan N., Biener G., Paprocki J., Thelen J.J., Raicu V., Zhao D. (2017): Two SERK receptor-like kinases interact with ems1 to control anther cell fate determination. Plant Physiology, 173: 326–337.

Li X., Salman A., Guo C., Yu J., Cao S., Gao X., Li W., Li H., Guo Y. (2018): 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, 7. doi: 10.3390/cells7090120

Liu Z., Wu Y., Yang F., Zhang Y., Chen S., Xie Q., Tian X., Zhou J.M. (2013): BIK1 interacts with PEPRs to mediate ethylene-induced immunity. Proceedings of the National Academy of Sciences of the United States of America, 110: 6205–6210.

Liu P.L., Xie L.L., Li P.W., Mao J.F., Liu H., Gao S.M., Shi P.H., Gong J.Q. (2016): Duplication and divergence of leucine-rich repeat receptor-like protein kinase (*LRR-RLK*) genes in basal Angiosperm *Amborella trichopoda*. Frontiers in Plant Science, 7. doi: 10.3389/fpls.2016.01952

Liu P.L., Du L., Huang Y., Gao S.M., Yu M. (2017): Origin and diversification of leucine-rich repeat receptor-like protein kinase (*LRR-RLK*) genes in plants. BMC Evolutionary Biology, 17. doi: 10.1186/s12862-017-0891-5

Lozano-Elena F., Caño-Delgado A.I. (2019): Emerging roles of vascular brassinosteroid receptors of the BRI1-like family. Current Opinion in Plant Biology, 51: 105–113.

Lu X., Shi H., Ou Y., Cui Y., Chang J., Peng L., Gou X., He K., Li J. (2020): RGF1-RGI1, a peptide-receptor complex, regulates *Arabidopsis* root meristem development via a MAPK signaling cascade. Molecular Plant, 13: 1594–1607.

Luu D.D., Joe A., Chen Y., Parys K., Bahar O., Pruitt R., Chan L.J.G., Petzold C.J., Long K., Adamchak C., Stewart V., Belkhadir Y., Ronald P.C. (2019): Biosynthesis and secretion of the microbial sulfated peptide RaxX and binding to the rice Xa21 immune receptor. Proceedings of the National Academy of Sciences of the United States of America, 116: 8525–8534.

Macho A.P., Zipfel C. (2014): Plant PRRs and the activation of innate immune signaling. Molecular Cell, 54: 263–272.

Magalhães D.M., Scholte L.L., Silva N.V., Oliveira G.C., Zipfel C., Takita M.A., De Souza A.A. (2016): LRR-RLK family from two *Citrus* species: Genome-wide identification and evolutionary aspects. BMC Genomics, 17. doi: 10.1186/s12864-016-2930-9

Meng J., Yang J., Peng M., Liu X., He H. (2020): Genome-wide characterization, evolution, and expression analysis of the leucine-rich repeat receptor-like protein kinase (LRR-RLK) gene family in *Medicago truncatula*. Life (Basel), 10. doi: 10.3390/life10090176

Mishra D., Suri G.S., Kaur G., Tiwari M. (2021): Comprehensive analysis of structural, functional, and evolutionary dynamics of Leucine Rich Repeats-RLKs in *Thinopyrum elongatum*. International Journal of Biological Macromolecules, 183: 513–527.

Miyazawa H., Oka-Kira E., Sato N., Takahashi H., Wu G.J., Sato S., Hayashi M., Betsuyaku S., Nakazono M., Tabata S., Harada K., Sawa S., Fukuda H., Kawaguchi M. (2010): The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. Development, 137: 4317–4325.

Mott G.A., Thakur S., Smakowska E., Wang P.W., Belkhadir Y., Desveaux D., Guttman D.S. (2016): Genomic screens identify a new phytobacterial microbe-associated molecular pattern and the cognate *Arabidopsis* receptor-like kinase that mediates its immune elicitation. Genome Biology, 17. doi: 10.1186/s13059-016-0955-7

Pertea M., Pertea G.M., Antonescu C.M., Chang T.C., Mendell J.T., Salzberg S.L. (2015): StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology, 33: 290–295.

Robinson M.D., McCarthy D.J., Smyth G.K. (2010): edgeR: A bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26: 139–140.

Shiu S.H., Bleecker A.B. (2001): Plant receptor-like kinase gene family: Diversity, function, and signaling. Science Signaling, 2001. doi: 10.1126/stke.2001.113.re22

Sun X., Wang G.L. (2011): Genome-wide identification, characterization and phylogenetic analysis of the rice LRR-kinases. PLoS One, 6. doi: 10.1371/journal.pone.0016079

Sun Y., Li L., Macho A.P., Han Z., Hu Z., Zipfel C., Zhou J.M., Chai J. (2013): Structural basis for flg22-induced activation of the *Arabidopsis* FLS2-BAK1 immune complex. Science, 342: 624–628.

https://doi.org/10.17221/131/2021-PPS

Sun J., Li L., Wang P., Zhang S., Wu J. (2017): Genome-wide characterization, evolution, and expression analysis of the leucine-rich repeat receptor-like protein kinase (LRR-RLK) gene family in Rosaceae genomes. BMC Genomics, 18. doi: 10.1186/s12864-017-4155-y

Tanaka K., Choi J., Cao Y., Stacey G. (2014): Extracellular ATP acts as a damage-associated molecular pattern (DAMP) signal in plants. Frontiers in Plant Scienc, 5. doi: 10.3389/fpls.2014.00446

Tiwari M., Pandey V., Singh B., Yadav M., Bhatia S. (2021): Evolutionary and expression dynamics of LRR-RLKs and functional establishment of KLAVIER homolog in shoot mediated regulation of AON in chickpea symbiosis. Genomics, 113: 4313–4326.

Wang Y., Tang H., Debarry J.D., Tan X., Li J., Wang X., Lee T.H., Jin H., Marler B., Guo H., Kissinger J.C., Paterson A.H. (2012): *MCScanX*: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Research, 40. doi: 10.1093/nar/gkr1293

Wang L., Albert M., Einig E., Fürst U., Krust D., Felix G. (2016): The pattern-recognition receptor CORE of Solanaceae detects bacterial cold-shock protein. Nature Plants, 2. doi: 10.1038/nplants.2016.185

Wang J., Liu S., Li C., Wang T., Zhang P., Chen K. (2017): PnLRR-RLK27, a novel leucine-rich repeats receptor-like protein kinase from the Antarctic moss *Pohlia nutans*, positively regulates salinity and oxidation-stress tolerance. PLoS One, 12. doi: 10.1371/journal.pone.0172869

Wang H., Chen Y., Wu X., Long Z., Sun C., Wang H., Wang S., Birch P.R.J., Tian Z. (2018): A potato Strubbelig-Receptor family member, StLRPK1, associates with StSERK3A/BAK1 and activates immunity. Journal of Experimental Botany, 69: 5573–5586.

Wang J., Hu T., Wang W., Hu H., Wei Q., Bao C. (2019): Investigation of evolutionary and expressional relationships in the function of the leucine-rich repeat receptor-like protein kinase gene family (LRR-RLK) in the radish (*Raphanus sativus* L.). Scientific Reports, 9. doi: 10.1038/s41598-019-43516-9

Wu Y., Xun Q., Guo Y., Zhang J., Cheng K., Shi T., He K., Hou S., Gou X., Li J. (2016): Genome-wide expression pattern analyses of the arabidopsis leucine-rich repeat receptor-like kinases. Molecular Plant, 9: 289–300.

Yan S., Ning K., Wang Z., Liu X., Zhong Y., Ding L., Zi H., Cheng Z., Li X., Shan H., Lv Q., Luo L., Liu R., Yan L., Zhou Z., Lucas W.J., Zhang X. (2020): CsIVP functions in vasculature development and downy mildew resistance in cucumber. PLoS Biology, 18. doi: 10.1371/journal.pbio.3000671

Yang D.L., Shi Z., Bao Y., Yan J., Yang Z., Yu H., Li Y., Gou M., Wang S., Zou B., Xu D., Ma Z., Kim J., Hua J. (2017): Calcium pumps and interacting BON1 protein modulate calcium signature, stomatal closure, and plant immunity. Plant Physiology, 175: 424–437.

Yeh Y.H., Panzeri D., Kadota Y., Huang Y.C., Huang P.Y., Tao C.N., Roux M., Chien H.C., Chin T.C., Chu P.W., Zipfel C., Zimmerli L. (2016): The *Arabidopsis* malectin-like/LRR-RLK IOS1 is critical for BAK1-dependent and BAK1-independent pattern-triggered immunity. Plant Cell, 28: 1701–1721.

Yu X., Feng B., He P., Shan L. (2017): From chaos to harmony: Responses and signaling upon microbial pattern recognition. Annual Review of Phytopathology, 55: 109–137.

Yuan N., Rai K.M., Balasubramanian V.K., Upadhyay S.K., Luo H., Mendu V. (2018): Genome-wide identification and characterization of LRR-RLKs reveal functional conservation of the SIF subfamily in cotton (*Gossypium hirsutum*). BMC Plant Biololy, 18. doi: 10.1186/s12870-018-1395-1

Zan Y., Ji Y., Zhang Y., Yang S., Song Y., Wang J. (2013): Genome-wide identification, characterization and expression analysis of populus leucine-rich repeat receptor-like protein kinase genes. BMC Genomics, 14. doi: 10.1186/1471-2164-14-318

Zhou F., Guo Y., Qiu L.J. (2016): Genome-wide identification and evolutionary analysis of leucine-rich repeat receptor-like protein kinase genes in soybean. BMC Plant Biology, 16. doi: 10.1186/s12870-016-0744-1

Zhu F., Deng J., Chen H., Liu P., Zheng L., Ye Q., Li R., Brault M., Wen J., Frugier F., Dong J., Wang T. (2020): A CEP peptide receptor-like kinase regulates auxin biosynthesis and ethylene signaling to coordinate root growth and symbiotic nodulation in *Medicago truncatula*. Plant Cell, 32: 2855–2877.

Zipfel C. (2008): Pattern-recognition receptors in plant innate immunity. Current Opinion in Immunology, 20: 10–16.

Zipfel C., Kunze G., Chinchilla D., Caniard A., Jones J.D., Boller T., Felix G. (2006): Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. Cell, 125: 749–760.

Received: September 28, 2021
Accepted: January 14, 2022
Published online: January 27, 2022