

Characterization, evolution, and abiotic stress responses of leucine-rich repeat receptor- like protein kinases in *Liriodendron chinense*

Zhiying Mu

Zhejiang A&F University

Mingyue Xu

Co-Innovation Center for Sustainable Forestry in Southern China, Nanning Forestry University

Teja Manda

Co-Innovation Center for Sustainable Forestry in Southern China, Nanning Forestry University

Jinhui Chen

Co-Innovation Center for Sustainable Forestry in Southern China, Nanning Forestry University

Liming Yang

yangliming@njfu.edu.cn

Co-Innovation Center for Sustainable Forestry in Southern China, Nanning Forestry University

Delight Hwarari

Co-Innovation Center for Sustainable Forestry in Southern China, Nanning Forestry University

Research Article

Keywords: *Liriodendron chinense*, LRR-RLK genes, gene expression, abiotic stress responses, phylogeny

Posted Date: February 1st, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-3905452/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on July 31st, 2024. See the published version at <https://doi.org/10.1186/s12864-024-10560-3>.

Abstract

Background

The *Liriodendron chinense* similar to all other flowering plants is vulnerable to plant extinction due to the severity of the abiotic stresses in the wake of the global climate change. Thereby, affecting its growth and development, and geographical distribution. Nonetheless, the *Liriodendron chinense* plays an essential role in the socio-economic and ecological sectors. The LRR-RLK genes comprise one of the largest groups of receptor-like kinases in plants, crucial for plant development and stress regulation; and the LRR-RLK genes have not been elucidated in the *Liriodendron chinense*.

Result

In this study, we identified 232 LchiLRR-RLK genes that were unevenly distributed on 17 chromosomes and 24 contigs, of which 67 paralogous gene pairs portrayed gene linkages; that contributed to LchiLRR-RLK gene family expansion through tandem (35.82%) and segmental (64.18%) duplications. Additionally, the synonymous and nonsynonymous ratios showed that the LchiLRR-RLK genes underwent a purifying or stabilizing selection during the evolutionary process. Investigation in the protein structures and domain conservation exhibited that LchiLRR-RLK carried conserved PK and LRR domains that also promoted their clustering in different subfamilies implicating gene evolutionary conservation. A deeper analysis of LchiLRR-RLK full protein sequences phylogeny showed 13 families that had a common ancestor protein. Interspecies gene collinearity showed more orthologous gene pairs between *L. chinense* and *P. trichocarpa*, suggesting various similar biological functions between the two plant species. Analysis of the functional roles of the LchiLRR-RLK genes using the qPCR demonstrated that they are involved in abiotic stress regulation, especially, members of subfamilies VIII, III, and Xa.

Conclusion

Conclusively, the LRR-RLK genes are conserved in the *L. chinense* and function to regulate the temperature and salt stresses, and this research provides new insights into understanding LchiLRR-RLK genes and their regulatory effects in abiotic stresses.

Introduction

The Leucine-rich repeat-like protein kinases (LRR-RLKs) play a crucial role in plant development and stress responses [1], and encompass one of the largest groups of receptor-like kinases in plants [2]. Their structures comprise of ectodomains and cytoplasmic domains, frequently occurring in a combination of LRR and RLK domains [3]. Their biochemical structure is composed of three functional domains: an extracellular domain (ECD) that perceives signals, a transmembrane domain that acts as an anchor to the protein within the membrane, and an intracellular kinase domain (KD) that transduces

signals downstream through autophosphorylation, followed by phosphorylation of exclusive substrates, subsequently [4]. In addition, the LRR-RLK ECD is characterized by varying numbers of LRR repeats that facilitate the sensing of a variety of ligands, including small molecules, peptides, or even entire proteins [5]. On the other hand, the LRR-RLK KD is common in protein kinases, constituting 12 conserved subdomains that exhibit an identical three-dimensional catalytic primary two-lobed structure; known for their vital functions in enzymatic roles [6, 7]. Furthermore, a typical plant LRR-RLK family is classified into 15–29 groups and subgroups based on the phylogenetic analysis of the KDs and denoted based on the subgroup classification of *A. thaliana* LRR-RLKs, numbered in Roman numerals [8]. The classification of the LRR-RLKs largely depends on the phylogeny of KDs due to the ambiguity in other conserved domains [9]. Nonetheless, the LRRs share a common structure defined by a 20–28 residue expanse rich in leucine, and seven discrete subfamilies have been identified sharing a conserved LxxLxLxxNxL(s/t)GxLPxxLxx (where L denotes the hydrophobic amino acid, and N stands for Asparagine, threonine, serine or cysteine, and x is a variable residue) [10]. In addition, the highly conserved region 'LxxLxLxxN' in LRRs conforms to a curved parallel β -sheet lining in the inner circumference of their solenoid structure, while the conserved 'L(s/t)GxLP' region forms the plant-specific β -strand which affects the positioning of the LRR stacks into a superhelical assemblage [11].

Extensive research has characterized the LRR-RLK gene family in several plants and showed their responses in the regulation of a wide range of biological processes in plants, such as growth and development, microsporogenesis and embryogenesis, plant immune response against pathogens, and tolerances to various abiotic stresses like heat, cold, drought, salt, and nutrient treatment [12, 13]. For instance, an LRR-RLK protein HSL3 has been shown to negatively regulate the stomatal closure by modulating the level of H_2O_2 in guard cells, thereby the HSL3 was concluded to participate in the regulation response of drought stress [14]. Overexpressed novel cold tolerance LRR-RLK gene (*MtCTLK1-OE*) in *M. truncatula* increased cold tolerance through activating the C-repeat-Binding Factor (CBF)-pathway, antioxidant defense system, and proline accumulation [15]. A phytosulfokine receptor (PSKR) in rice was upregulated by the ABA, and increased the stomatal closure, regulated the ROS activity in the guard cells, thereby enhancing drought tolerance in Arabidopsis. Further research demonstrated that the OsPSRR15 enhanced drought stress through direct interaction with AtPYL9 and its orthologue OsPYL11 through its kinase domain in the plasma membrane and nucleus [16]. In other studies, the somatic embryogenesis receptor kinase (SERK), an LRR-RLK gene in wheat was shown to perform a dual function in heat tolerance and biotic stress resistance to *P. striform* F. sp. *triticii* through interacting with the TaDJA7 to activate the HTSPs [17]. Previous research has also shown an LRR-RLK gene, Phloem Intercalated with Xylem-Like 1 (PXL1) in Arabidopsis to be induced by cold and heat stress, and phosphorylate AtHIRD1 and AtLHCA1 in the regulation of the signal transduction pathways under temperature fluctuations [18].

The above-mentioned studies provide substantial evidence of the involvement of the LRR-RLK genes in the regulation of various environmental stresses and their active participation during plant growth and development. Additionally, several publications have identified the LRR-RLK genes in Arabidopsis [19], *Zea mays* [20], rosacea plants [21], *Saccharum* [22], *Gossypium* species [23], and others. However, the

LRR-RLKs have not yet been identified and their abiotic stress response elucidated in *Liriodendron chinense* (Lchi). The *Liriodendron* genus comprises two prominent species, the *Liriodendron chinense* and *Liriodendron tulipifera* [24]. The plant species vary from annual plants on account of their woody secondary growth and perennial habit [25]. Similarly, several abiotic stresses affect the *Liriodendron chinense* growth and distribution, including cold, heat, drought, light, and nutrient utilization stresses [26]. To gain insight into the response of LRR-RLK genes in *L. chinense*, we investigated their bioinformatic properties and expression patterns to various abiotic stresses through qPCR expression analysis. Although these investigations are not exhaustive, they provide a firm foundation for further biological experimentation.

Methods

Identification and Classification of LRR-RLK Genes

The genomic and protein sequences of *Liriodendron chinense* were obtained from the TreeGene database (<https://treegenesdb.org/>, accessed on 30 March 2023); those of *Arabidopsis thaliana* were retrieved from the from TAIR (<https://www.arabidopsis.org/browse/genefamily/leuc.jsp>; accessed on the 30th of March 2023) and used as a reference in the identification process. Other LRR-RLK genes from other plants were obtained from the Phytozome v13 (<https://phytozome-next.jgi.doe.gov/>) including *P. patens*, *s. moellendorffii*, and *O. sativa*. To identify LRR-RLK genes in the *Liriodendron chinense*, we used the method by [21]. Putative PKs were initially obtained by searching the Hidden Markov Models of the typical Pkinase clade [Pkinase (PF00069) and Pkinase_Tyr (PF07714)] obtained from the Pfam database v.28 [27], against the proteome of *L. chinense* using the simple HMMER search in Tootools [28]; with an E-value cut-off of 0.0001. Typical PKs were identified with coverage of the Pfam domain model of at least 50%, after the screening, and CDD search from both the NCBI CDD (<https://ncbi.nlm.nih-gov.brum.beds.ac.uk/Structure/bwrpsb/bwrpsb.cgi>) and SMART (<http://smart.embl-heidelberg.de/>) was used to further authenticate the identified putative LchiLRR-RLKs. In addition, previously defined HMMs of different typical PKs families and subfamilies (https://github.com/lileiting/Plant_Pkinase_fam.hmm) [29], were used to classify the identified PKs into families and subfamilies at an E-value cut-off of 0.0001. The HMM subfamily was classified based on the phylogenetic classification of model plant species, *A. thaliana*, *O. sativa*, *P. patens*, and *C. reinhardtii* [29]. To confirm this classification, a maximum likelihood phylogenetic tree was constructed with full-length protein sequences using the Mega X software at a 100 bootstrap value. Multiple alignments were done using the MUSCLE in Mega X. The tree was beautified using the online ITOL tool [30].

Multiple Alignments and Phylogenetic Analysis

Multiple sequence alignments were performed on the full-length amino acid sequences of LRR-RLK proteins in *L. chinense*, *A. thaliana*, *O. sativa*, and *P. patens* with the MUSCLE program using default parameters as implemented in MEGA. Subsequently, MEGA X software [31] was used to construct a phylogenetic tree based on the alignments using the neighbor-joining tree (NJ) method. The bootstrap

test was replicated 100 times using the p-distance model. To confirm the result from the NJT method, the phylogenetic tree was constructed using the Maximum likelihood (ML) method. All identified LRR-RLK proteins were predicted for subcellular localization using the DeepLoc-2.0 tool (<https://services.healthtech.dtu.dk/services/DeepLoc-2.0/>; accessed on the 27th of April 2023). The conserved domain (CDD), motif number and arrangements, and cis-elements were analyzed using previous methods.

Chromosome location, Gene Duplication, and Synteny analysis.

The chromosome location of each LRR-RLK gene was obtained from their genome resources. Then the Tbtools software was used to map the gene on the corresponding chromosome. For synteny analysis, genome regions that showed syntenic relationships were identified using the McScanX in Tbtools with default parameters. The synonymous and non-synonymous ratios (Ka/Ks) were calculated using the Ka/Ks calculator in Tbtools. The tandem and segmental repeated genes were searched by comparing the LRR-RLK gene in their corresponding positions in chromosome/scaffolds, and adjacent genes were designated as tandem duplicated genes.

RNA isolation, cDNA synthesis, and the RT-qPCR analysis of the LRR-RLK gene in *L. chinense*

Total RNA was extracted from the tender leaves using the HiScript® III 1st Strand cDNA Synthesis Kit (+ gDNA wiper) (Nanjing Vazyme Biotech Co., Ltd; China) according to the manufacturer's instructions. RNA degradation and contamination were monitored on 1% agarose gel, and concentration was measured using the Nanophotometer spectrophotometer (IMPLEN, CA, USA). Samples of total RNA (1 µg) were used to synthesize cDNA, the cDNA was subsequently diluted to 100 ng/µL and used as a template for RT-qPCR analysis. The experiment was set in three treatments, cold, heat, and salt stresses; analyzed at four time points (0h, 3h, 24h, and 3 days (3d)), and Actin was used as the reference gene. Primers used in this experiment were designed with Snap Gene Viewer (<https://www.snapgene.com/>), and are shown in the Supplementary Data Table S1. The relative gene expression levels were determined using the 2- $\Delta\Delta C_t$ method [32].

Results

Genome-wide identification and Classification of the LRR-RLK genes in *Liriodendron chinense*.

We searched the annotated genes in the *Liriodendron chinense* genome resource for putative PKs and identified 1488 typical PKs (Supplementary File 1). After removing redundant, overlapping, and sequences lacking the LRR-RLK conserved domains, a total of 232 LchiLRR-RLK protein sequences remained. The obtained LchiLRR-RLK protein sequences carried an extracellular domain (ECD), a transmembrane domain, and an intracellular kinase domain (KD) (Fig. 1a). In addition, the LchiLRR-RLK ECD was branded by varying numbers of LRR. The obtained PKs were then classified and renamed into families and subfamilies based on the previous classifications in Arabidopsis and rice model plants (Supplementary File 2) [4]. LchiLRR-RLK protein classification showed 13 families, which were named in

Roman numerals (I-XV) following previous publications. Interestingly, families XIV, IX, VII and other subfamilies present in *A. thaliana* were absent in *L. chinense*. Of the present families VIII and I were the largest subfamilies with 65 and 61 members, respectively. The other families consisted of not more than 25 members, of which Xa had the least members, with only 2 sequences. Furthermore, the protein sequence lengths varied, ranging between 100 and 1454 and the isoelectric point of the obtained LchiLRR-RLK protein sequences ranged from 4.77 to 10.41 (Supplementary File 2). Suggesting that the obtained LchiLRR-RLKs ranged from weakly acid to strong basic. All the proteins showed a cellular localization in the plasma membrane (Supplementary File 2).

Gene chromosomal location, duplications, and collinearity

To obtain further insights on the gene locations of the identified LchiLRR-RLK protein sequences, the TBtools software was used to map each gene location on the chromosome and contig (Fig. 1b; Supplementary File 2). Results showed that the LchiLRR-RLKs were unevenly distributed on 17 chromosomes and 24 contigs, each chromosome carried at least 12 and at most 27 LchiLRR-RLK genes. Additional information on the LchiLRR-RLKs is listed in the Additional file. Additional analysis of the obtained genes' collinearity exhibited a total of 67 paralogous gene pairs, constituting almost half of the total number (Fig. 1b; Supplementary File 3). This finding further suggested gene duplication events and gene family expansion within the LchiLRR-RLK gene family. Therefore, to understand the mode of gene expansion, we compared two main gene duplication events: tandem and segmental duplication events [33, 34] (Fig. 1b; Supplementary File 3). Results showed that of the total 67 duplicated gene pairs, 24 pairs were tandem arrays, contributing 35.82% (48/134) of the duplicated genes. In addition, most of the tandem duplications were obtained in the VIII and XII families contributing 67% (16/24) of the total tandem arrays. On the other hand, 43 gene pairs were by segmental duplications and contributed 64.18% (86/136) to the expansion of the LRR-RLK gene family expansion in *L. chinense* (Supplementary File 3). Most of the segmental duplications were obtained in the I and XI families contributing 23% (20/86) and 12% (10/86) of the total number, similar results were obtained in Rosaceae plant genomes [21].

The synonymous and nonsynonymous values and their ratios are used to estimate the selection pressure of a given protein or DNA experience. To have an insight into the source of duplicate genes, we calculated the synonymous and nonsynonymous values and their ratios (Supplementary File 3). A total of 138 linked genes were obtained, of which all the genes investigated for substitution mutation had a K_a/K_s ratio less than 1 ($K_a/K_s < 1$), signifying a purifying or stabilizing selection of the LchiLRR-RLK genes during the evolutionary process.

Motif, Gene Structure, and Domain Conservation analyses reveal conserved evolution.

To gain insight into the 232 LchiLRR-RLK protein functions, we computed the conserved motif numbers and arrangement (Fig. 2; Supplementary Fig. 1; Supplementary Fig. 2). We observed that the investigated proteins clustered based on their similar motif arrangements and possible phylogenetic relationships (Fig. 2a). In detail, most groups had at most 9 motifs present in each protein except for groups III and XII

which had 10 to 15 motifs (Fig. 2b). In addition, we noticed that motifs 9 and 10 among others were to a greater extent present in the groups III, XI, and XII only; while motifs 5 and 2 were abundant in all groups, suggesting their full conservation.

Generally, the basic structure of the LRR-RLK gene comprises a PK domain and an LRR domain. We investigated the conserved domain in identified gene candidates. In this study, we showed that different subfamilies have different compositions of protein domains (Fig. 2c). However, the majority of the subfamilies carried a Pkinase domain and an LRRNT_2 (leucine-rich repeat N-terminal) (Fig. 2b; Supplementary File 1; Supplementary Fig. 1). We also observed that some subfamilies like the XI, II, and XIII had a mixture of both the Pkinase and the Pkinase_Tyr. The subfamily XI had the most conserved domains, carrying at most 6 conserved domains. Gene structure prediction is vital in comprehending the gene evolution of a gene family [26]. In this study, we analyzed the gene structures of 232 LchiLRR-RLK genes (Fig. 2d; Supplementary Fig. 1). Results showed that exon and intron numbers varied with gene sequences. Families Xb-2, Xa-1, VII-2, XII, XI-1, III, IX, and Xb-2 had exon ranges between 1–3, accompanied by 2 or 1 introns flanking the N- and C- terminals. While the rest of the gene families had numbers more than 3. In addition, gene family XIIIb had the greatest number of exons up to 30, without introns. Interestingly, the exon structures differed among the gene families, gene families with fewer exon numbers carried elongated exons, and intron structures were almost of similar sizes among the different gene families, except for the gene family Xb-3 which was flanked by an elongated intron in the C-terminal

Phylogenetic analysis of the LRR-RLK gene family

Systematic classification of a gene family based on the protein phylogeny facilitates the building of functional and genomic studies. In this study, 1032 LRR-RLK full protein sequences from five plant species, *L. chinense*, *A. thaliana*, *O. sativa*, *S. moellendorffii*, and *P. patens*, were used to construct a phylogenetic tree using the neighbor-joining tree (NJT) method in MEGA X (Fig. 3a). Results displayed a clustering of protein sequences into various families and subfamilies consistent with previous publications. We observed a total of 19 cluster groups that had diverged from 3 main branches, forming 14 families and 5 subfamilies. Nonetheless, all the cluster groups were observed to have diverged from a common ancestral protein. A deeper analysis showed that one of the main branches carried most of the LRR-RLK groups, 11 in total, while the remaining had 1 and 6 groups. This fact suggests that the LRR-RLKs evolved mainly from a single ancestral protein that diversified possibly through speciation adaptation and other evolutionary measures. Comparisons of the protein quantities in various cluster groups showed that groups XIII and VIII had the most protein numbers, 148 and 187 protein sequences, respectively (Fig. 3b). In contrast, groups XV and II had the least number of protein sequences, 9 and 14, respectively. Additionally, some groups lacked representation of other plant species, for instance, the *L. chinense* LRR-RLKs were absent in groups, XI, XIIIb, Xa, VI, and XIV. On the other hand, some groups were fully represented, suggesting similarity in functionality of LRR-RLKs from different plant species. In total, this finding suggests different degrees of gene expansions within LRR-RLK groups and that proteins clustered together may exhibit similar structures and functions.

Plant Synteny

Gene collinearity within different plant species genes may also reflect phylogenetic relations and possibly similar gene functions. In this research, we used the TBtools software to compute LRR-RLK gene collinearity between four plant species: *A. thaliana*, *L. chinense*, *O. sativa*, and *P. trichocarpa* (Fig. 3c). Results showed a dense linkage of several genes. Specifically, *L. chinense* had 55, 48, and 89 orthologous gene pairs with *A. thaliana*, *O. sativa*, and *P. trichocarpa*, respectively. Suggesting a closer evolutionary relationship between *L. chinense* and *P. trichocarpa* than any other plant, although recent research has suggested that *O. sativa* is evolutionarily closer to *L. chinense*. In addition, this result may suggest more similar biological functions between *P. trichocarpa* and *L. chinense*.

Cis-regulatory Elements

Evaluation of the cis-regulatory elements present in the promoter region regions is critical in the understanding of transcriptional regulation and gene function [35]. In this study, a 1.5kb region of the identified *LchiLRR-RLK* genes was considered a potential promoter region. A total of 5056 putative elements were identified and categorized into three response factors, growth and development, plant phytohormone, and biotic and abiotic responses using the Plant Care Online Database (Supplementary Fig. 1, Additional File 4). However, 24 representatives are shown in the manuscript for presentation purposes (Fig. 4). The cis-element abundancies were not consistent in all the 232 *LchiLRR-RLKs*. Nonetheless, we noted that distributions of the cis-elements followed a somewhat similar pattern within identical families and subfamilies.

Comparisons in the above-mentioned response factors showed an overrepresentation of the cis-elements in the biotic and abiotic responses constituting 55.22% of the total identified cis-elements. Suggesting that the *LchiLRR-RLK* genes are more invested in biotic and abiotic stress functional roles. Furthermore, the phytohormonal and growth and development responses constituted 16.8% and 28%, respectively. In-depth analysis showed that the *LchiLRR-RLK* families VIII and I had the most cis-element in all the response factors analyzed probably due to the fact they have many members present as compared to other families. Particularly, this research focused more on the abiotic stresses; therefore our analysis also exhibited several cis-regulatory elements involved in the abiotic stress responses including the DRE-core, LTR, STRE, MYB., etc. Specifically, the MYB cis-elements were present in all the *LchiLRR-RLK* genes, while the STRE and LTR elements were also present in almost all the *LchiLRR-RLKs*. Suggesting that LRR-RLKs in *L. chinense* respond to abiotic stresses including temperature and drought stresses.

Figure 4. Cis-regulatory element analysis. The total numbers of the identified putative cis-elements in the promoter of *L. chinense* LRR-RLK genes are shown in the boxes. Different colors show the total ranges of cis-elements present.

Protein Interaction and Protein Analyses

Protein-to-protein interaction (PPI) analysis is crucial in elucidating protein function and the impact of protein absence or presence. In this study, we investigated the protein interaction between various LRR-RLK proteins using the Online String database (Fig. 5a). We observed that the protein families were densely interconnected. Individual LRR-RLK gene groups interacted with various families probably for efficient biological functions. Suggesting that the LRR-RLK gene families in *L. chinense* interact for full protein function. In detail, most proteins were linked with the Lchi_IV-1 of group IV showing a possibility that

Lchi_IV-1 acts as a control hub group mediating several protein functions. Previously, Chen et al. [11] have shown that plants with numerous continuous LRRs and few insertion segments in the ectodomain tend to stack into super helical shapes for sensing various ligands in signal activations [11].

To gain insight into the protein structures of LchiLRR-RLK proteins, we searched for the homology models using the SWISS-model online tool (Fig. 5b). Previous research has established that LRR assembly structures are predictable due to the high conservation of the LRR repeats, with the “LxxLxLxxN” forming the inner side of the superhelix, while the “xLs/tG” form the plant-specific second β -sheet on the lateral side, and the remainder forming the backside [12, 20]. In this study, ten representative LchiLRR-RLK proteins showed different protein structures however, those from groups XI, XII, XIII, and XV portrayed a somewhat similar structure. Generally, the LchiLRR-RLKs had numerous LRRs that formed the superhelices and buried their hydrophobic patches inside (Fig. 5). Additionally, the conserved residues of the LRR backbone were more hydrophobic than the variable residues, nonetheless, the variable residues had lower hydrophilicity than we predicted to aid in the protein proper folding.

qPCR Expression Analysis

To understand the possible responses of the LchiLRR-RLK genes to three abiotic stresses, twenty LchiLRR-RLK genes were selected for qPCR analysis; and their expression patterns in response to cold, heat, and salt stresses were analyzed over three time points, 3hr, 24h, and 3 days; and compared against the control (0h) (Fig. 6). Generally, most of the analyzed LchiLRR-RLK genes showed significant gene expression trends as compared to the control (0h). Particularly, in cold stress (Fig. 6a), LchiLRR-RLKs clustered into five notable expression groups that exhibited a high to low-expression trend; of which most of the genes had a high upregulation. Noteworthy, three genes, Lchi_I-24/Xa-3/II-1 showed the highest upregulation from stress onset to termination as compared to the 0h. Interestingly, Lchi_VIII-61 clustered within the same group although it exhibited an alternating expression trend, which was marked by a downregulation at 3h, upregulation at 24h, and finally a downregulation at 3d. Nonetheless, this expression trend does not have any significant differences when compared to the 0h, signifying that these four genes respond positively to cold stress. The second and third gene cluster groups had seven LchiLRR-RLK genes that had extremely low expression trends marked by downregulations from stress onset to termination, suggesting that these genes respond negatively to cold stress. The remaining groups had a fairly upregulated expression, although some genes also exhibited an alternating expression trend. Concluding that these genes also respond positively to cold stress to a certain extent.

Similarly, in heat stress, the LchiLRR-RLK genes grouped into clusters according to their expression trends (Fig. 6b). Like cold stress, the first group genes including, Lchi_Xa-3, and others were characterized by the highest positive expression patterns. These exhibited upregulations from stress onset to termination at 3d as compared to the 0h. Suggesting that these four genes are highly responsive to cold stress. Likewise, the second and third groups had fairly upregulated expression patterns, also characterized by alternating gene regulation trends that had upregulations at 3h and downregulations at 24h and an upregulation at stress termination. Surprisingly, LchiVIII-61 and Lchi_XI(2)-8 had extremely low expression values at 3d. Lastly, the fourth and fifth groups had extremely low expression patterns from stress onset up to termination, however, Lchi_I-28 in the fifth group was highly expressed at 3d. Probably due to the fact it responds to heat stress after a long time of exposure.

In salt stress, LchiLRR-RLK genes clustered according to their similarities in response to stress (Fig. 6c). Generally, most of the analyzed genes had an extremely down-regulated expression at 3d, with a few in the second cluster group exhibiting an upregulation at 3d. These included, Lch_Xa-3/VIII-63/I-32/I-28, and others. This result can be concluded as thus, most LchiLRR-RLK genes are downregulated during the salt stress, and a few respond positively to the cold stress at long periods of exposure to stress. In summary, the abiotic stress analysis demonstrated that the LchiLRR-RLK genes respond to cold, heat, and salt stresses at varying extents, suggesting that these genes may regulate the abiotic stresses at different time points.

Discussion

The LRR- RLK genes constitute one of the largest gene families in plants, playing a major role in plant growth and development, and biotic and abiotic responses [1]. In addition, various RLK genes have been elucidated, including the pathogenesis-related protein 5-like receptor kinase (PR5K), epidermal growth factor-like repeats (EGF), lectin-binding domain (LB), tumor necrosis factor receptor-like (TNFR), and the S-domain [12, 36]. The 232 identified LRR-RLK genes in *L. chinense* carried an extracellular domain (ECD), a transmembrane domain, and an intracellular kinase domain (KD) with the ECD branded by varying numbers of LRR repeats. Additional analysis revealed 15 LRR motifs with a 24 residue-long LRR domain, L/cxxLxxNxL/fsGxI/1PxxL/lxx (Fig. 1), this was in agreement with the previous finding of a plant LRR denoted by a LxxLxxLxLxxNxLxGxIPxxLxx consensus sequence [37].

Investigations in the CDD and motif analyses also exhibited a conserved PK domain and an LRR domain. These findings demonstrated that the LRR-RLK genes are conserved in the *L. chinense* and may be involved in different functions. Adams et al. [7] have shown that protein kinases are known for their vital functions in enzymatic roles due to the presence of conserved subdomains [7]. In addition, the LRR-RLK genes have been identified in several plant species, including *Populus trichocarpa*, citrus species, Rosaceae species, maize, and others. In this study, we identified 232 LRR-RLK genes, which were far more than in *A. thaliana* and *O. sativa*. In rice, this can be accounted for by the fact that *L. chinense* has a large genome size of 749.3kbp compared to *A. thaliana* and *O. sativa* with 374. Kbp [38]. Also, the inconsistencies in the gene family sizes can be related to gene duplication events. Research has related

gene family expansion mainly due to two duplication events, tandem, and segmental duplication as sources of gene family expansion as it increases gene and genome densities [39]. In this study, we showed that both the tandem and segmental duplications contributed 16% and 79% of the gene expansion of duplicated genes in the *L. chinense* LRR-RLK gene family. Similarly, previous studies in the Rosaceae gene families have shown that tandem and segmental duplications are two major forms of gene family expansion contributing to almost 50% of the total gene family expansions [21]. In-depth analysis showed that individual families and subfamilies were consistently expanded through tandem duplication. Interestingly, the families that had expanded through the tandem duplications also had the greatest numbers of the LRR-RLKs. Other research has also established that the expansion of the LRR-RLK gene family is enhanced due to their prime function in both development and defense responses, and continuous selection pressure imposed by the development complexities in the environment; reflecting LRR-RLK random gene drift [37]. Previous studies have also shown that the expansion of the LRR-RLK gene family has contributed to RLK genes through both adaptive and non-adaptive evolution [29].

The origin of the LRR-RLK gene family remains a mystery nonetheless, research has shown that the domain shuffling of the LRR and KD has led to the founding of the RLK subfamilies [2]. To understand the phylogenetic relationships among the LRR-RLKs, we computed the phylogenetic tree using LRR-RLK full proteins from five plant species. The phylogenetic classification of the LRR-RLK proteins in *L. chinense* was similar to previous publications [36, 40]. The protein sequences clustered into 21 families and subfamilies (I - XV) based on similar protein and domain arrangements. The groups XIII and VIII had the most proteins, suggesting that protein duplication was relatively high in these families. Generally, the LchiLRR-RLK evolution showed a divergence into several groups emanating from an ancestral LchiLRR-RLK gene. Based on tree topologies, we also observed that some phylogenetic groups lacked representation from other plant species. For instance, *L. chinense* LRR-RLK proteins were absent in groups, XI, XIIb, VI, and XIV while present in *A. thaliana* and the lower plants had a full representation, this may entail that these genes in *A. thaliana* and the lower plants diverged recently, or were lost in *L. chinense* during angiosperm WGD duplication events [41]. Another reason for gene group absence may be related to group function specialization, for example, the PRK in subfamily II and PSY in subfamily XI were established in early plants due to their specific function in the pollen tube development [42, 43]. While other families are present in flowering plants due to that they control flowering in vascular plants [44]. In agreement with this finding, Liu et al. [1] have further published that subfamilies I and VII-2 evolved from a common ancestor before the divergence of specific lineages and that most LRR-RLK subfamilies were established in land plants before the divergence of moss [1].

However other LchiLRR-RLKs clustered within the same families with other LRR-RLKs from other plant species such as *A. thaliana*; suggesting that these proteins may exhibit similar functional roles to their paralogues. Furthermore, some LchiLRR-RLKs clustered in similar groups with LRR-RLKs from *P. patens* and *S. moellendorffi*, research has marked the mosses and lycophytes as early forms of plant life [45]. Suggesting that LchiLRR-RLKs are well conserved and little function loss has been experienced due to

gene mutations and related processes; and that they possess central roles in the regulation of common developmental and defense pathways of different land plant lineages [41].

Plant LRR-RLKs are important membrane-localized receptors sensing various ligands to regulate plant developmental processes. Their diversity allows for response to several environmental stresses and actively functions in growth and developmental processes. In other studies, the somatic embryogenesis receptor kinase (SERK), an LRR-RLK gene in wheat was shown to perform a dual function in heat tolerance and biotic stress resistance in *P. striiformis* F. sp. *tritici* through interacting with the TaDJA7 to activate the HTSPs [17]. Previous research has also shown an LRR-RLK gene, Phloem Intercalated with Xylem-Like 1 (PXL1) in Arabidopsis to be induced by cold and heat stress, and phosphorylate AtHIRD1 and AtLHCA1 in the regulation of the signal transduction pathways under temperature fluctuations [18]. In this research, we investigated the cis-regulatory elements present in the promoter regions of identified LRR-RLK genes. Our results showed that the LRR-RLKs are actively involved in growth and development, and biotic and abiotic stress responses. In detail, the biotic and abiotic response elements constituted a total of 55.22% of the total identified cis-elements. The identified cis-regulatory elements included the DRE-core, LTR, STRE, MYB, WRE3, and the WUN-motif. Suggesting that LchiLRR-RLKs are invested in abiotic stress regulation, especially with the abundance of the Low-Temperature Response elements and the DRE. To further elucidate the expression patterns of the observed LchiLRR-RLK genes, the qPCR expression analysis revealed that a large proportion of the identified genes responded to both the temperature and salt stresses. In detail, most members of groups VIII, II, and Xa LchiLRR-RLKs had the highest expression patterns showing that these families regulate the abiotic stresses. A recent study in *Medicago truncatula* has shown that the MtCTLK1 an LRR-RLK gene increased cold tolerance through inducing the expression of the CBFs and CBF-dependent cold responsive genes. Further analysis in this research also evidence that MtCTLK1 increased antioxidant enzyme activities and proline accumulation [15]. Providing possible insights that the LchiLRR-RLKs can also regulate the cold stress linking the CBF-cold response pathway [46]. Another LRR-RLK protein HSL3 was shown to negatively regulate stomatal closure by modulating the level of H₂O₂ in guard cells, thereby regulating drought and salt stress [14]. Various RLKs from different subfamilies, such as RPK1, CYSTEINE-RICH RLK (CRK36), PROLINE-RICH-EXTENSIN-LIKE RLK4 (PREK4), and the GUARD CELL HYDROGEN PEROXIDE-RESISTANT 1 (GHR1) in Arabidopsis have also been reported to also regulate salt stress [47–52].

Conclusion

The LRR-RLK genes mediate a multiple of signal transduction pathways, thereby are involved in several plant processes including the regulation of abiotic stresses. In this study, we performed in silico analyses of the LRR-RLK gene in *Liriodendron chinense* that enabled the identification of 232 LchiLRR-RLK genes localized on 17 chromosomes and 24 contigs. Analysis of their physiochemical properties through the protein motif numbers and arrangements, conserved domain, and gene structures exhibited that LRR-RLK proteins cluster together in different subfamilies depending on similarity and conservation. Evolutionary studies demonstrated that these subfamilies have a shared evolution history that indicates molecular function [53]. A deeper survey into the LchiLRR-RLK genes promoter sequences evidenced

that they carry cis-regulatory elements that respond to abiotic stresses including the low-temperature stress. Using RNA-seq data and qPCR expression, we also concluded that a great number of LchiLRR-RLK genes may regulate the heat, cold, and salt stress especially members of the subfamilies VIII and III. Although these results are not exhaustive, they provide a basis for future molecular experiments.

Abbreviations

LRR-RLK: Leucine-rich repeat-like protein kinases; Lchi: *Liriodendron chinense*; ECD: extracellular domain; KD: intracellular kinase domain; *MtCTLK1-OE*: *M. truncatula* cold tolerance LRR-RLK- Overexpressed; CBF: C-repeat-Binding Factor; PSKR: phytosulfokine receptor; SERK: somatic embryogenesis receptor kinase; PKs: Protein Kinases; HMM: Hidden Markov Models; CDD: Conserved domain; ML: Maximum likelihood; Ka/Ks: synonymous and non-synonymous ratios; 2- $\Delta\Delta$ Ct: 2 (-delta delta CT); DRE-core: Dehydration responsive elements-core; LTR: low-temperature responsive elements; PPI: Protein-to-protein interaction; qPCR; quantitative polymerase chain reaction; WGD; whole genome duplication.

Declarations

Acknowledgments

We thank the funders of this research for their financial support.

Data Availability Statement

Genome and gene model annotations files are available on the NCBI website (https://www.ncbi.nlm.nih.gov/assembly/GCA_003013855.2, accessed on 4 June 2023).

Supplementary Files

The following supporting information can be downloaded at..

Author Contribution List

Y.L. and H.D. conceived, planned, coordinated the project, and finalized the manuscript. Z.M., M.X., Y.G., and H.D. performed the experiments, and data analysis and wrote the draft. T.M. and J.C. validated and contributed to data analysis and curation, and revised the manuscript. L.Y. coordinated, contributed to data curation, and finalized, and funded this research.

Funding

This research work was funded by the National Natural Science Foundation of China (No. 31971682, 32071784), the Research Startup Fund for High-Level and Highly-Educated Talents of Nanjing Forestry University.

Competing Interests

The authors declare no conflict of interest.

References

1. Liu P-L, et al. Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evol Biol.* 2017;17(1):47.
2. Shiu S-H, Bleecker AB. *Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases.* Proceedings of the National Academy of Sciences, 2001. 98(19): p. 10763–10768.
3. Park S, et al. Molecular dissection of the response of a rice leucine-rich repeat receptor-like kinase (LRR-RLK) gene to abiotic stresses. *J Plant Physiol.* 2014;171(17):1645–53.
4. Gou X, et al. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in *Arabidopsis thaliana*. *BMC Genomics.* 2010;11(1):1–15.
5. Bojar D, et al. Crystal structures of the phosphorylated BRI 1 kinase domain and implications for brassinosteroid signal initiation. *Plant J.* 2014;78(1):31–43.
6. Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science.* 1988;241(4861):42–52.
7. Adams JA. Kinetic and catalytic mechanisms of protein kinases. *Chem Rev.* 2001;101(8):2271–90.
8. Shiu S-H, et al. Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell.* 2004;16(5):1220–34.
9. Han GZ. Origin and evolution of the plant immune system. *New Phytol.* 2019;222(1):70–83.
10. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol.* 2001;11(6):725–32.
11. Chen T. Identification and characterization of the LRR repeats in plant LRR-RLKs. *BMC Mol Cell Biology.* 2021;22(1):9.
12. Tyagi S, Shumayla, Upadhyay SK. Chap. 13 - *Role of leucine-rich repeat receptor-like kinases in abiotic and biotic stress responses in plants.* Plant Receptor-Like Kinases. Academic Press; 2023. pp. 239–55. S.K. Upadhyay and Shumayla, Editors.
13. Soltabayeva A, et al. Receptor-like Kinases (LRR-RLKs) in response of plants to biotic and abiotic stresses. *Plants.* 2022;11(19):2660.
14. Liu X-s, et al. The LRR-RLK protein HSL3 regulates stomatal closure and the drought stress response by modulating hydrogen peroxide homeostasis. *Front Plant Sci.* 2020;11:548034.
15. Geng B, et al. A novel LRR-RLK (CTLK) confers cold tolerance through regulation on the C-repeat-binding factor pathway, antioxidants, and proline accumulation. *Plant J.* 2021;108(6):1679–89.
16. Nagar P, et al. OsPSKR15, a phytosulfokine receptor from rice enhances abscisic acid response and drought stress tolerance. *Physiol Plant.* 2022;174(1):e13569.
17. Shi Y et al. *The LRR-RLK Protein TaSERK1 Positively Regulates High-temperature Seedling Plant Resistance to Puccinia striiformis f. sp. tritici through Interacting with TaDJA7.* Phytopathology,

2023(ja).

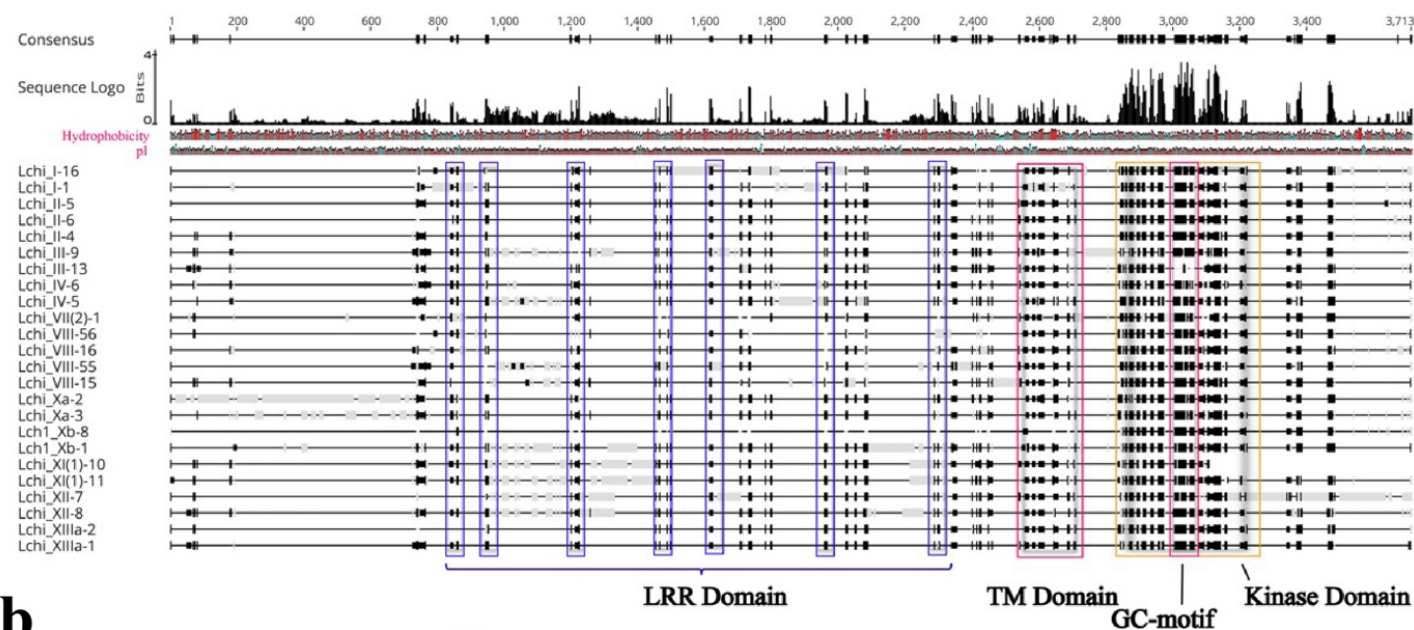
18. Jung CG, et al. Molecular characterization of the cold-and heat-induced Arabidopsis PXL1 gene and its potential role in transduction pathways under temperature fluctuations. *J Plant Physiol.* 2015;176:138–46.
19. Kemmerling B, et al. A genome-wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant immunity. *Front Plant Sci.* 2011;2:88.
20. Song W et al. *Identification of immune related LRR-containing genes in maize (Zea mays L.) by genome-wide sequence analysis.* International journal of genomics, 2015. 2015.
21. Sun J, et al. 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.* 2017;18(1):1–15.
22. Cheng W, et al. Genome-Wide identification of LRR-RLK family in saccharum and expression analysis in response to biotic and abiotic stress. *Curr Issues Mol Biol.* 2021;43(3):1632–51.
23. Sun R, et al. Genome-wide analysis of LRR-RLK gene family in four Gossypium species and expression analysis during cotton development and stress responses. *Genes.* 2018;9(12):592.
24. Cao Y, et al. Alterations in Population Distribution of Liriodendron chinense (Hemsl.) Sarg. and Liriodendron tulipifera Linn. Caused by Climate Change. *Forests.* 2022;13(3):488.
25. Hwarari D, et al. Systematic Characterization of GATA Transcription Factors in Liriodendron chinense and Functional Validation in Abiotic Stresses. *Plants.* 2023;12(12):2349.
26. Hwarari D, et al. Comprehensive Bioinformatics and Expression Analysis of TCP Transcription Factors in Liriodendron chinense Reveals Putative Abiotic Stress Regulatory Roles. *Forests.* 2022;13(9):1401.
27. Mistry J, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* 2020;49(D1):D412–9.
28. Chen C, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13(8):1194–202.
29. Lehti-Shiu MD, Shiu S-H. Diversity, classification and function of the plant protein kinase superfamily. *Philosophical Trans Royal Soc B: Biol Sci.* 2012;367(1602):2619–39.
30. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49(W1):W293–6.
31. Kumar S, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35(6):1547.
32. Chen Z, Halford NG, Liu C. Real-Time Quantitative PCR: Primer Design, Reference Gene Selection, Calculations and Statistics. *Metabolites.* 2023;13(7):806.
33. Zhu B. *Tandem Duplications, Segmental Duplications and Deletions, and Their Applications.* in *Computer Science–Theory and Applications: 15th International Computer Science Symposium in Russia, CSR 2020, Yekaterinburg, Russia, June 29–July 3, 2020, Proceedings 15.* 2020. Springer.

34. Freeling M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu Rev Plant Biol.* 2009;60:433–53.
35. Chang W-C, et al. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics.* 2008;9(1):1–14.
36. Liu P-L, et al. Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evol Biol.* 2017;17(1):1–16.
37. Albrecht C, et al. Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE proteins serve brassinosteroid-dependent and-independent signaling pathways. *Plant Physiol.* 2008;148(1):611–9.
38. Ouyang S, et al. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res.* 2006;35(suppl1):D883–7.
39. Cantsilieris S, et al. An evolutionary driver of interspersed segmental duplications in primates. *Genome Biol.* 2020;21(1):1–35.
40. Zan Y, et al. Genome-wide identification, characterization and expression analysis of populus leucine-rich repeat receptor-like protein kinase genes. *BMC Genomics.* 2013;14(1):318.
41. Soltis PS, Soltis DE. Ancient WGD events as drivers of key innovations in angiosperms. *Curr Opin Plant Biol.* 2016;30:159–65.
42. Chang F, et al. AtPRK2 promotes ROP1 activation via RopGEFs in the control of polarized pollen tube growth. *Mol Plant.* 2013;6(4):1187–201.
43. Fisher K, Turner S. PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. *Curr Biol.* 2007;17(12):1061–6.
44. Floyd SK, Bowman JL. The ancestral developmental tool kit of land plants. *Int J Plant Sci.* 2007;168(1):1–35.
45. Anderberg HI, Kjellbom P, Johanson U. Annotation of *Selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants. *Front Plant Sci.* 2012;3:33.
46. Hwarari D, et al. ICE-CBF-COR signaling cascade and its regulation in plants responding to cold stress. *Int J Mol Sci.* 2022;23(3):1549.
47. Tanaka H, et al. Abiotic stress-inducible receptor-like kinases negatively control ABA signaling in *Arabidopsis*. *Plant J.* 2012;70(4):599–613.
48. Hua D, et al. A plasma membrane receptor kinase, GHR1, mediates abscisic acid-and hydrogen peroxide-regulated stomatal movement in *Arabidopsis*. *Plant Cell.* 2012;24(6):2546–61.
49. Lee IC, et al. Age-dependent action of an ABA-inducible receptor kinase, RPK1, as a positive regulator of senescence in *Arabidopsis* leaves. *Plant Cell Physiol.* 2011;52(4):651–62.
50. Osakabe K, Osakabe Y, Toki S. *Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases*. *Proceedings of the National Academy of Sciences*, 2010. 107(26): p. 12034–12039.
51. Bai L, et al. Plasma membrane-associated proline-rich extensin-like receptor kinase 4, a novel regulator of Ca²⁺ signalling, is required for abscisic acid responses in *Arabidopsis thaliana*. *Plant J.*

52. Osakabe Y, et al. Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis. Plant Cell. 2005;17(4):1105–19.
53. Kileeg Z et al. *Differential expansion and retention patterns of LRR-RLK genes across plant evolution*. bioRxiv, 2023: p. 2023.07.26.549740.

Figures

a



b

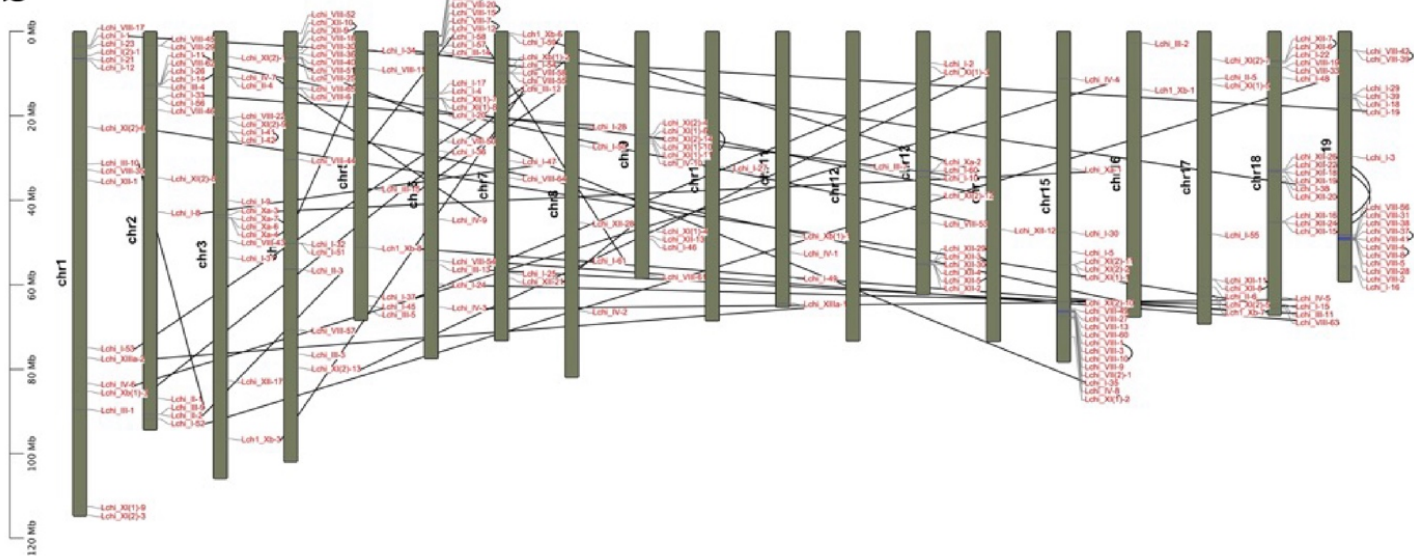


Figure 1

LRR-RLK gene characterization in *L. chinense*. **a** shows the multiple sequence alignments (MSA) of the LchiLRR-RLK representative protein sequence in each group as generated by the Geneious Prime

boundaries. **b** The interspecies phylogenetic tree, generated from selected plant orthologs by the Xshell software (<https://xshell.en.softonic.com/>, accessed on 23 June 2023) and the online tool iTOL [30]. Below is the summary of the total number of LRR-RLK proteins present in each plant and group in Figure 3A. The background heat map shows varying group sizes of LRR-RLKs in each plant analyzed. **c** shows interspecies LRR-RLK gene collinearity between *L. chinense*, *A. thaliana*, *O. sativa*, and *P. trichocarpa*. Blue curvy lines show collinear GATA genes between the four plant species, each chromosome was named above.

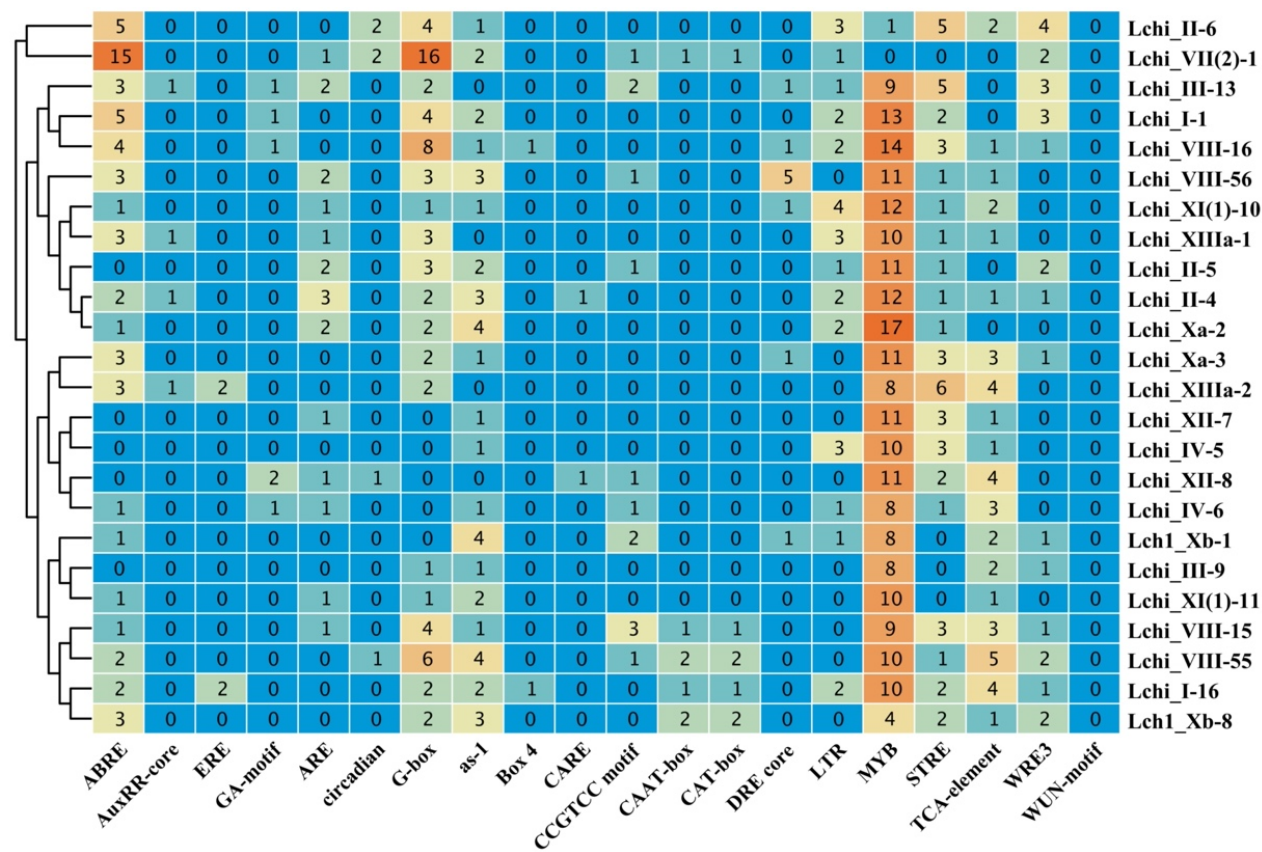


Figure 4

Cis-regulatory element analysis. The total numbers of the identified putative cis-elements in the promoter of *L. chinense* LRR-RLK genes are shown in the boxes. Different colors show the total ranges of cis-elements present.

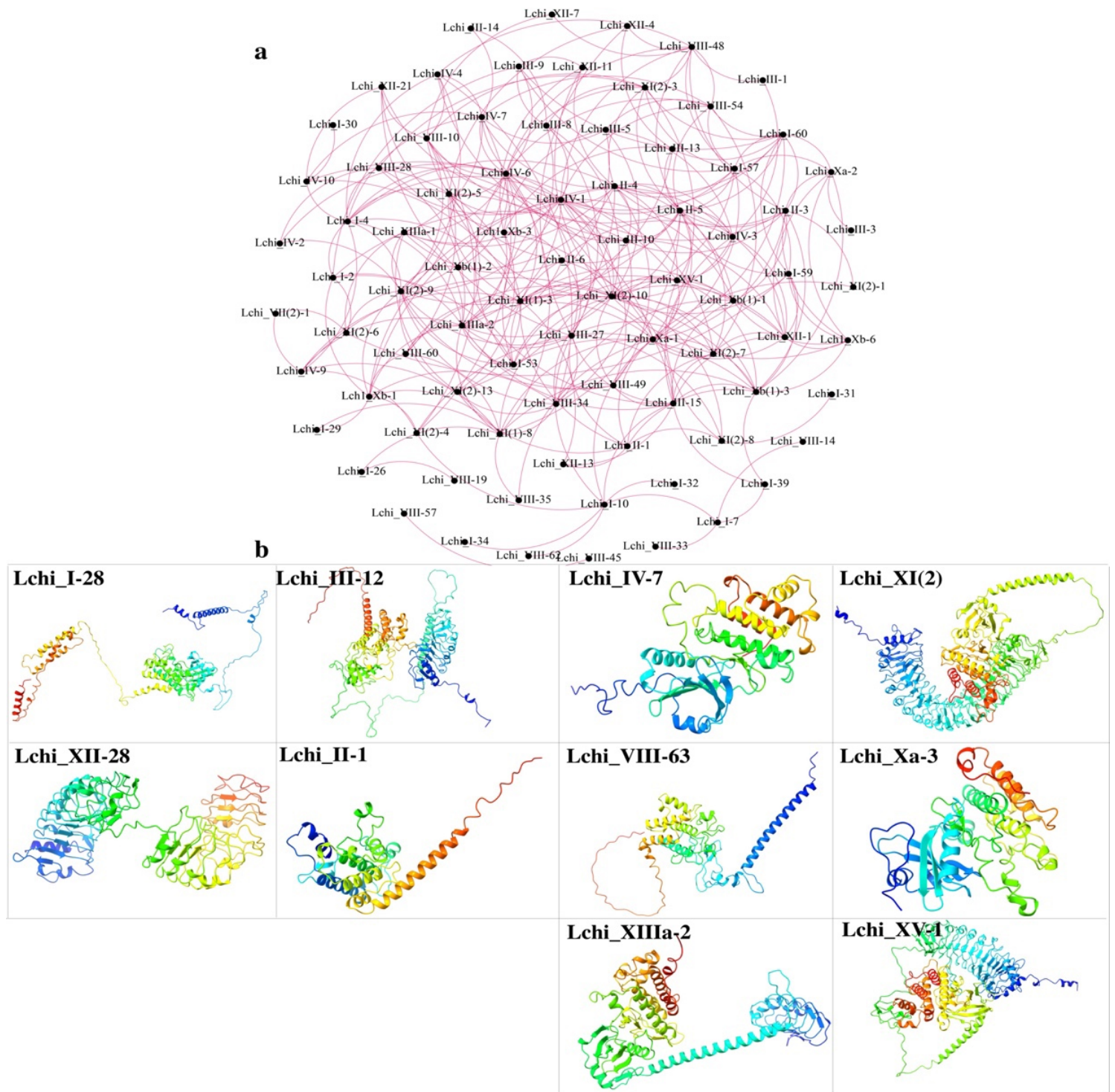


Figure 5

Protein interaction and structure analysis. **a** Protein-protein interaction network for LRR-RLKs was analyzed using the STRING website (<http://string-db.org>, accessed on 25 June 2023) using the full-length protein sequences of the LRR-RLK family. *Arabidopsis thaliana* was used as a reference plant species. Each LRR-RLK protein is labeled at the node, and the red line depicts interactions. **b** The 3D protein structure prediction of 10 LRR-RLK proteins, showing the potential strands, helices, and coil formation.

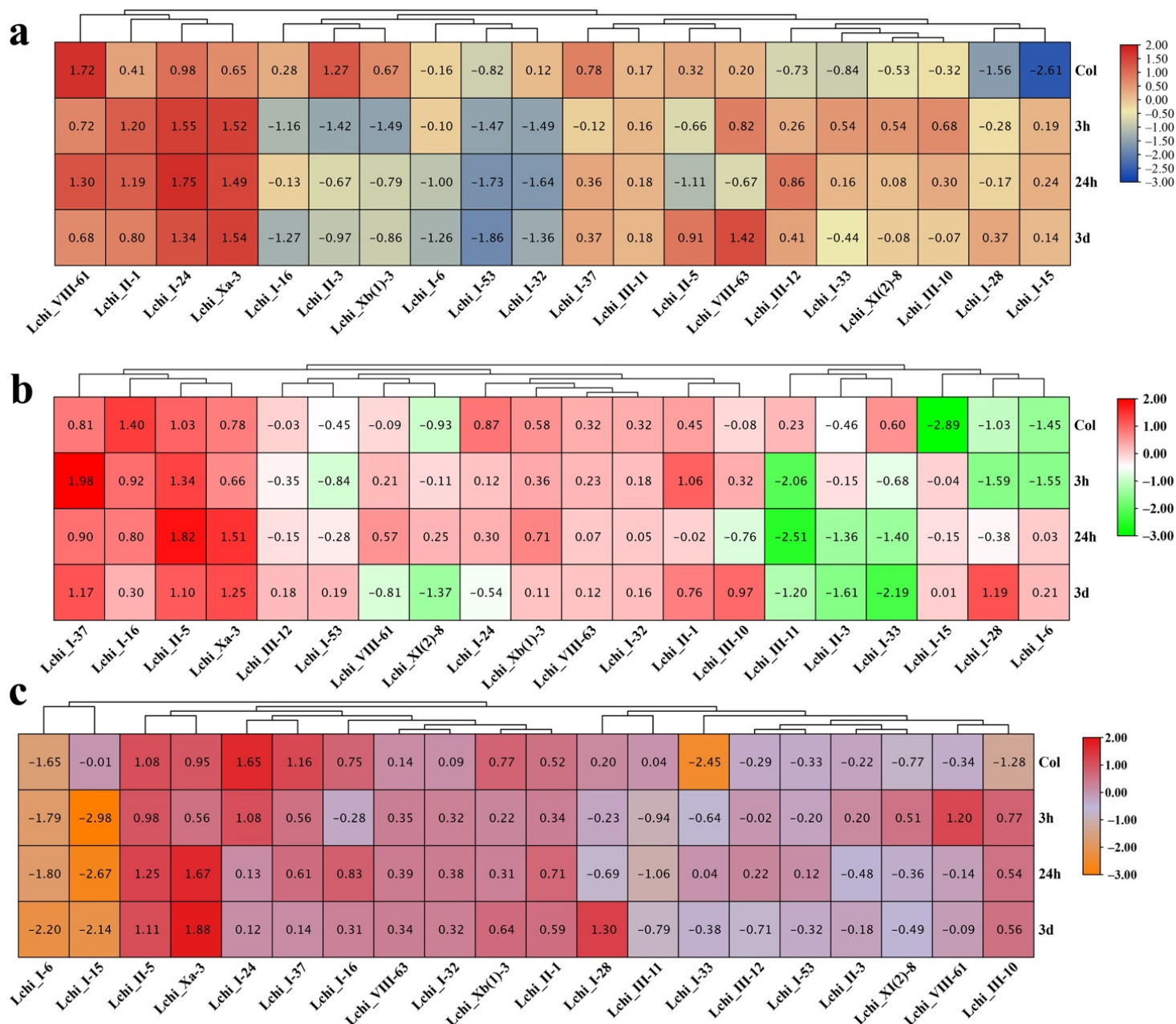


Figure 6

qPCR gene expression analysis of 20 LchiLRR-RLK genes in different abiotic stresses (cold, heat, and salt) at different time points (3h, 24 h, and 3 d) generated using the TBtools software. Expression was analyzed based on the relative mRNA levels, which were reduced using the log value and maximized to 1. The final values were illustrated as heatmaps, with the key in the right upper corner. **a** cold, **b** heat, and **c** salt stresses.

Supplementary Files

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

- [SupplementaryFile.zip](#)