Canadian Science Publishing | Genome

In silico genomic characterization, classification, and expression profiling of the LRR-RLK family in Lactuca sativa under stress and developmental conditions

Journal:	*Genome*
Manuscript ID	gen-2025-0085.R1
Manuscript Type:	Research Article
Date Submitted by the Author:	09-Jan-2026
Complete List of Authors:	Pereira, Welison Andrade; Federal University of Lavras Department of Biology, Departamento de Biologia Oliveira, Elenildo dos Santos; Federal University of Lavras Department of Biology, Departamento de Biologia Pereira, Kelly Aparecida Da Cunha ; Federal University of Lavras Department of Biology, Departamento de Biologia
Is the manuscript for consideration in a Special Issue or Collection?:	Not applicable (regular submission)
Keyword:	Lettuce, genome-wide, LRR-RLK, gene expression, transcriptome

In silico genomic characterization, classification, and expression profiling of the LRR-RLK family in *Lactuca sativa* under stress and developmental conditions

Welison Andrade Pereira[1]*, Elenildo dos Santos Oliveira[1], Kelly Aparecida da Cunha Pereira[1]

[1]Laboratório de Genética Molecular, Departamento de Biologia, Universidade Federal de Lavras (UFLA), Lavras, Minas Gerais, Brasil.

*Corresponding author: Welison Andrade Pereira – E-mail: welison.pereira@ufla.br

Abstract

Leucine-rich repeat receptor-like kinases (LRR-RLKs) constitute the largest receptor kinase family in plants and are involved in diverse processes related to development, environmental perception, and signal transduction. Despite their relevance, the LRR-RLK gene family has not been systematically characterized in *Lactuca sativa*. In this study, a genome-wide identification and integrative characterization of *LsLRR-RLK* genes was performed using curated bioinformatic pipelines. Phylogenetic relationships, protein architecture, conserved motifs, biochemical properties, gene structure, chromosomal distribution, duplication patterns, syntenic relationships, cis-regulatory elements, and transcriptomic profiles were comprehensively analyzed. A total of 269 high-confidence LRR-RLK genes were identified and classified into established subfamilies, revealing marked subfamily-specific patterns of structural organization, regulatory architecture, and transcriptional behavior. Expression analyses across developmental stages and multiple environmental conditions demonstrated widespread but heterogeneous transcriptional regulation within the family. Collectively, this work provides a detailed genomic and functional overview of the LRR-RLK gene family in *L. sativa*, establishing a structured reference for comparative and exploratory studies of receptor-like kinases in plants.

Key words: Lettuce, genome-wide, LRR-RLK, gene expression, transcriptome

1. Introduction

Receptor-like kinase (RLK) proteins generally comprise a signal peptide, an extracellular domain (ECD) often containing leucine-rich repeats (LRRs), a single pass transmembrane domain (TM), and an intracellular kinase domain (KD) (Shiu and Bleecker 2001a, 2001b; Gou et al. 2010; Liu et al. 2017). This structural configuration confers broad biological versatility to LRR-RLKs, enabling their participation in essential processes such as growth, development, signal transduction, and responses to biotic and abiotic stress (Zan et al. 2013; Yuan et al. 2018; Zhiqi et al. 2024). Given their central role in signal perception and transduction, LRR-RLKs are considered promising targets for improving crop resilience, which is increasingly relevant in the context of global climate change (Zan et al. 2013; Sun et al. 2018).

LRR-RLKs represent the largest RLK subfamily in plants (Shiu and Bleecker 2001b; Lehti-Shiu et al. 2009; Zan et al. 2013; Zhou et al. 2016; Magalhães et al. 2016; Sun et al. 2017; Yuan et al. 2018; Cheng et al. 2021), playing key roles in adaptation and survival under diverse environmental conditions (Zan et al. 2013; Wu et al. 2016; Yuan et al. 2018). These proteins can function as pattern recognition receptors (PRRs), detecting pathogen-associated molecular patterns (PAMPs), such as flagellin (recognized by FLS2) (Gómez-Gómez and Boller 2000; Sun et al. 2013) and EF-Tu (recognized by EFR) (Zipfel et al. 2006). LRR-RLKs also regulate developmental processes including meristem maintenance (CLV1), somatic embryogenesis and brassinosteroid signaling (SERKs) (Yu et al. 2023), anther development (BAM1–3) (Hord et al. 2006), vascular differentiation (Caño-Delgado et al. 2004), and hormonal pathways such as abscisic acid signaling (Qin et al. 2016). This broad functional spectrum underscores their importance in integrating growth, development, and environmental responsiveness (Sun et al. 2018).

While LRR-RLK families have been characterized in several species, including *Arabidopsis thaliana* (Shiu and Bleecker 2003), *Oryza sativa* (Shiu et al. 2004), *Solanum*

lycopersicum (Wei et al. 2015), *Glycine max* (Zhou et al. 2016), *Gossypium hirsutum* (Sun et al. 2018), *Medicago truncatula* (Meng et al. 2020), *Saccharum spontaneum* (Cheng et al. 2021), *Phaseolus vulgaris* (da Silva Dambroz et al. 2023), *Liriodendron chinense* (Mu et al. 2024), among others (Yang et al. 2024; Zhang et al. 2024; Jiang et al. 2024), comprehensive genome-wide analyses of this family are still lacking in *Lactuca sativa*, a crop of major nutritional and economic importance. Addressing this gap is important for improving our understanding of the diversity and potential roles of LRR-RLKs in lettuce and for informing future breeding strategies.

In this study, we conducted a comprehensive *in silico* genomic, structural, and expression-based characterization of the *LRR-RLK* gene family in *L. sativa*. We identified and classified *LRR-RLK* genes based on the presence of kinase, transmembrane, and LRR domains, leveraging HMMs of LRR-RLK subfamilies derived from 25 plant species (Lehti-Shiu and Shiu 2012). Phylogenetic and physicochemical prediction were performed to explore the structural motifs, subcellular localization, and domain organization of the LsLRR-RLK proteins. Additionally, we examined synteny with *A. thaliana* and *Solanum lycopersicum* to identify conserved genomic regions that may suggest shared functional roles or evolutionary origins. Furthermore, we investigated the cis-regulatory elements in promoter regions and analyzed the expression patterns of *LsLRR-RLK* genes across various tissues and under different biotic and abiotic stress conditions using transcriptomic data. Overall, this study outlines the genomic landscape of LRR-RLKs in *L. sativa*, providing a framework for future studies.

2. Materials and Methods

The methodological workflow of this study is summarized in figure 1.

2.1. Sequence retrieval and identification of LRR-RLKs

Protein sequences of *Lactuca sativa* (genome version 8) were retrieved from Phytozome v13 (Goodstein et al. 2012). Candidate LRR-RLKs were identified using InterProScan v5.76 (Jones et al. 2014), integrating Pfam, SMART, and CDD databases (Marchler-Bauer et al. 2017; Letunic et al. 2021; Mistry et al. 2021) to detect leucine-rich repeat (LRR) domains and serine/threonine protein kinase domains, applying an E-value threshold of $\leq 1 \times 10^{-5}$. Proteins were retained when they exhibited the canonical plant LRR-RLK architecture, defined by an extracellular LRR region, a single-pass transmembrane helix, and a C-terminal intracellular kinase domain (Lehti-Shiu and Shiu 2012; Man et al. 2020). Redundant or overlapping LRR predictions were collapsed following the merging criteria proposed by (Man et al. 2020).

Structural validation and topology-based curation were performed using a multi-tool pipeline combining signal peptide prediction and membrane topology inference (SignalP v6.0, TMHMM 2.0, TOPCONS 2.0, and CCTOP) (Krogh et al. 2001; Dobson et al. 2015; Tsirigos et al. 2015; Teufel et al. 2022). Domain architecture and membrane topology were visually inspected using Protter (Omasits et al. 2014). Sequences showing multi-pass membrane topology, inverted domain orientation, missing essential components, or incomplete kinase domains were excluded. For loci with multiple predicted isoforms, a single representative sequence was selected based on structural completeness and topological consistency and length. All inclusion and exclusion decisions are documented (Table S1).

2.2. Phylogenetic analysis

Phylogenetic analyses were conducted using full-length amino acid sequences of LRR-RLK proteins. Sequences were aligned with MAFFT v7 using the L-INS-i strategy (Katoh et al. 2018), and maximum-likelihood trees were inferred with IQ-TREE v2, employing ModelFinder for automatic substitution model selection and 1,000 ultrafast bootstrap replicates to assess branch support (Nguyen et al. 2015). Two phylogenetic reconstructions were generated using identical alignment and inference parameters: a combined analysis including

LRR-RLKs from *Lactuca sativa* and *Arabidopsis thaliana* (Shiu and Bleecker 2001a), rooted with an AGC_MAST serine/threonine kinase from *L. sativa* (Lsat_1_v5_gn_0_34000.1), and a species-specific analysis restricted to *L. sativa* LRR-RLKs.

Subfamily classification of LsLRR-RLKs was performed using an integrative framework combining Hidden Markov Model (HMM) profiling, based on subfamily-specific models defined by (Lehti-Shiu and Shiu 2012), and phylogenetic placement relative to well-characterized *A. thaliana* LRR-RLKs. HMM results were used as the primary criterion for subfamily assignment, while phylogenetic relationships were used to support classification and resolve ambiguous cases. Phylogenetic trees were visualized and annotated using iTOL v6 (Letunic and Bork 2024).

2.3. Protein domain identification and architectural curation

Protein domain composition and architecture were further characterized using Pfam and CDD profile models implemented in the NCBI Conserved Domain Database (Marchler-Bauer et al. 2017), applying an E-value threshold of ≤ 0.01. Domain predictions were interpreted according to the canonical structural organization of plant LRR-RLKs, comprising an extracellular LRR region, a single-pass transmembrane helix, and a C-terminal cytoplasmic kinase domain.

Overlapping predictions of the Pkinase (PF00069) and PK_Tyr/Ser-Thr (PF07714) domains were resolved by retaining the domain associated with the lowest E-value. Canonical LRR profiles (PF08263, PF13855, PF12799) were consolidated into continuous extracellular LRR blocks by selecting the best-supported prediction in overlapping regions while preserving non-overlapping LRR units. Domains other than LRR and kinase were retained only when they did not overlap with elements of the canonical LRR-RLK architecture, thereby minimizing redundancies arising from Pfam profile fragmentation or low-confidence predictions.

Transmembrane helices were predicted using TMHMM, CCTOP, and TOPCONS, with final assignments primarily based on consensus among methods. In cases of disagreement, hydrophobic segments of 16–30 amino acids located immediately downstream of the LRR region were prioritized, consistent with typical receptor-like kinase topology. Signal peptides were predicted using SignalP v6.0; when absent, N-terminal signal peptide predictions from CCTOP were accepted if their length (16–60 amino acids) was compatible with canonical signal peptides. The curated domain architectures were subsequently used for structural interpretation and visualization using the domain illustration tools implemented in TBtools (Chen et al. 2020).

2.4. Conserved motif analysis

Conserved structural motifs in LsLRR-RLK proteins were identified using MEME Suite v5.5.7 (Bailey et al. 2015) based on full-length amino acid sequences from the curated dataset. Motif discovery was performed using the zero or one occurrence per sequence (ZOOPS) model, with motif lengths constrained between 6 and 50 amino acids and a predefined maximum number of motifs, while other parameters were kept at default settings. Motif composition and distribution were examined across the entire LsLRR-RLK dataset, at the individual protein level, and among subfamilies. Motif positions and frequencies were visualized using TBtools.

2.5. Physicochemical properties and subcellular localization prediction of the LsLRR-RLKs

Physicochemical properties of LsLRR-RLK proteins, including protein length, molecular weight (MW), theoretical isoelectric point (pI), and grand average of hydropathicity (GRAVY), were calculated using the Sequence Manipulation Suite 2 - SMS2 (Stothard 2000).

Amino acid composition was determined for each protein and expressed as relative amino acid percentages. Compositional patterns across the LsLRR-RLK dataset were visualized using heatmaps generated with TBtools (Chen et al. 2020).

Subcellular localization was predicted using ProtComp (SoftBerry Inc.), CELLO v2.5, and WoLF PSORT (Yu et al. 2006; Horton et al. 2007), and the results were compiled to provide a comparative analysis of predicted cellular compartments of LsLRR-RLK proteins.

2.6. Gene structure, chromosomal localization, and SNP marker association

Exon-intron structures of *LsLRR-RLK* genes were analyzed using GSDS (Hu et al. 2015) based on genomic and coding sequences retrieved from Phytozome. Chromosomal positions of *LsLRR-RLK* genes were obtained from the same source and used to map their distribution across *Lactuca sativa* chromosomes using PhenoGram (Wolfe et al. 2013). SNP marker data previously reported for lettuce cultivars (Park et al. 2022) were integrated with *LsLRR-RLK* genomic coordinates to assess physical proximity between markers and gene loci. Gene-SNP associations were defined using a ±1 Mb genomic window, enabling the identification of genomic regions co-localizing with *LsLRR-RLK* genes for subsequent descriptive analyses.

2.7. Synteny analysis

Comparative synteny analyses were conducted to identify conserved genomic relationships involving *LsLRR-RLK* genes through comparisons with *Arabidopsis thaliana* and *Solanum lycopersicum*. Genome sequences, gene coordinates, and annotation files (GFF3) for all species were retrieved from Phytozome.

Syntenic gene pairs were identified using the One Step MCScanX workflow and processed using the Text Merge tools implemented in TBtools (Wang et al. 2012; Chen et al. 2020). Syntenic relationships involving *LsLRR-RLK* genes were extracted and visualized using the Multiple Synteny Plot module, focusing on conserved colinear gene pairs.

2.8. Gene duplication analysis

Gene duplication patterns of *LsLRR-RLK* genes were classified into tandem, proximal, segmental (including whole-genome duplication–derived), and dispersed categories using

MCScanX (Wang et al. 2012). The analysis was based on an all-versus-all BLASTP search of the *Lactuca sativa* proteome (E-value ≤ 1×10^{-5}), combined with gene positional information extracted from genome annotation (GFF3) files.

Duplication types were defined according to the standard MCScanX criteria, whereby tandem duplicates correspond to adjacent homologous genes on the same chromosome, proximal duplicates represent nearby homologs separated by a small number of intervening genes, segmental duplicates are paralogous gene pairs located within conserved collinear blocks, and dispersed duplicates lack evidence of proximity or collinearity. The distribution of duplication types was summarized at both whole-family and subfamily levels.

2.9. Ka/Ks analysis and selective pressure estimation

Selective pressure acting on duplicated *LsLRR-RLK* genes was evaluated by estimating nonsynonymous (Ka) and synonymous (Ks) substitution rates for paralogous gene pairs identified in the duplication analysis. Protein sequences were aligned using MAFFT v7, after which codon-based alignments were generated by back-translation to nucleotide sequences. Ka, Ks, and Ka/Ks (ω) values were estimated using the yn00 method implemented in the PAML package (Yang 2007). Gene pairs yielding unreliable estimates, including Ks = 0 or excessively large Ks values, were excluded from subsequent analyses, and Ka/Ks ratios were summarized according to duplication type and LRR-RLK subfamily.

2.10. Cis-regulatory elements

Promoter sequences corresponding to 2,000 bp upstream of the transcription start site (Meng et al. 2020; Cheng et al. 2021) of each *LsLRR-RLK* gene were retrieved from the *Lactuca sativa* genome assembly (Phytozome v13). Cis-regulatory elements (CREs) were identified using PlantCARE (Lescot 2002) and grouped into eight functional categories: abiotic stress-responsive, biotic stress-responsive, core promoter elements, development-related, hormone-

194 responsive, light-responsive, unknown, and unnamed, based on PlantCARE annotations and
195 curated literature. To avoid artificial inflation of CRE counts, overlapping motif predictions
196 were curated according to a non-overlapping rule, so that each nucleotide position contributed
197 to a single CRE occurrence. Overlapping matches of the same motif were collapsed into the
198 maximum number of non-overlapping instances per promoter. Promoter sequences containing
199 stretches of undefined nucleotides ("N") were retained in the analyses and explicitly indicated
200 in the results.

201 **2.11. Gene expression analysis under stress and developmental conditions**

202 Transcriptional responses of *LsLRR-RLK* genes were investigated using publicly
203 available RNA-Seq datasets encompassing abiotic stresses (nutrient deficiencies, temperature,
204 drought, and salinity), biotic interactions (fungal pathogens and mycorrhizal symbiosis), and
205 tissue-specific expression profiles (Verwaaijen et al. 2019; Leong et al. 2023; Pink et al. 2023;
206 Park et al. 2023; Krishnamoorthi et al. 2024; Shi et al. 2024; Vangelisti et al. 2024). Detailed
207 experimental designs and contrasts are provided in Table S2.

208 Raw reads from each study were processed independently using a standardized RNA-
209 Seq pipeline, including quality control, trimming, and alignment with FastQC, Trimmomatic,
210 and STAR (Andrews 2010; Dobin et al. 2013; Bolger et al. 2014). Reads were mapped to the
211 *Lactuca sativa* reference genome (Lsat_Salinas_v7) using gene models from Ensembl Plants,
212 and gene-level count matrices were generated for downstream analyses. Differential expression
213 analyses were performed using DESeq2 implemented via the iDEP platform, applying
214 significance thresholds of FDR < 0.05 and $|\log_2$ fold change$| \geq 1.5$. Analyses were strictly
215 restricted to within-study contrasts, avoiding direct quantitative comparisons across
216 independent datasets. Normalized expression values were extracted and visualized using
217 heatmaps and intersection analyses implemented in TBtools.

3. Results

3.1 Identification and structural curation of LRR-RLK genes in *Lactuca sativa*

An initial InterProScan survey identified 483 protein isoforms from 296 genomic loci containing LRR- and kinase-related signatures. Topology-based structural curation and isoform consolidation led to the exclusion of 214 isoforms, yielding a final dataset of 269 high-confidence LRR-RLK proteins. The retained proteins displayed the canonical type I receptor-like kinase architecture, consisting of an extracellular LRR region, a single transmembrane helix, and a C-terminal serine/threonine kinase domain. Excluded isoforms predominantly exhibited inconsistent membrane topology, inverted or incompatible domain organization, or incomplete LRR or kinase regions. All curation decisions were based on integrated topology predictions and manual inspection, with detailed justification provided in Table S1.

3.2. Phylogenetic relationships and subfamily classification of LsLRR-RLKs

Across both phylogenetic reconstructions, LsLRR-RLKs were classified into 23 subfamilies, exhibiting substantial variation in size and phylogenetic structure (Tables S3 and S4). Subfamily assignments inferred from phylogenetic placement were largely consistent with Hidden Markov Model (HMM)-based classification using subfamily-specific profiles derived from 25 plant species (Lehti-Shiu and Shiu 2012). In the combined *Lactuca sativa–Arabidopsis thaliana* phylogeny, 18 subfamilies formed monophyletic groups, whereas five (LRR-I-1, -XI-1, -III, -VII-3, and -Xb-2) showed polyphyletic patterns (Fig. S1). Subfamily sizes varied widely, with LRR-XI-1, -III, -I-1, -XII-1, and -II representing the largest groups. All major clades were supported by high bootstrap values, despite polyphyletic configurations in some large subfamilies (LRR-XI-1, -III, and -I-1).

In the *Lactuca sativa*–only phylogeny, the same 23 subfamilies were recovered (Fig. 2). In this dataset, 22 subfamilies formed monophyletic clades, whereas the LRR-XI-1 subfamily

was resolved into two distinct clades of comparable size (39 and 38 members), along with a single isolated member (Tables S3 and S4).

3.3. Protein domain identification and architectural curation

Sequence analysis confirmed that the identified LsLRR-RLK proteins display the canonical receptor-like kinase architecture, consisting of an extracellular LRR-rich region, a single transmembrane helix, and a C-terminal kinase domain (Fig. 3A-B).

Most proteins (98.1%) harbored a PK_Tyr_Ser-Thr–type kinase domain, whereas Pkinase domains were confined to members of subfamilies LRR-XI-1 and -III (Fig. 3B). Signal peptides were detected in 90.7% of the LsLRR-RLKs (Table S5). Consistent with previous reports for this gene family, structural variation was primarily associated with the extracellular region. Subfamilies differed markedly in both the number and composition of LRR repeats, with LRRNT_2, LRR_4, and LRR_8 representing the predominant profiles, present in 54.3%, 21.2%, and 84.8% of the proteins, respectively. Subfamily-specific patterns ranged from compact architectures with low LRR content (mean ≤2 repeats per protein) to expanded ectodomains containing more than five LRR repeats. Subfamilies LRR-VII-1, -VII-3, -XI-1, -XII-1, and -Xb-1 exhibited the highest LRR densities, reaching up to ten repeats, whereas LRR-VIII-2 and -Xa displayed more compact configurations.

Accessory domains beyond the canonical LRR-RLK framework were highlighted. Malectin and malectin-like domains were predominantly associated with the LRR-I-1 and -VIII-2 subfamilies, respectively, whereas the Island domain was exclusive to LRR-Xb-1. Other non-LRR domains occurred sporadically, with no evidence of subfamily-wide expansion.

3.4. Structural motifs

Conserved motif analysis identified 15 distinct motifs across the 269 curated LsLRR-RLKs and one AGC_MAST (Fig. 4A). Per-protein counts ranged from 7 to 35 (mean = 16.9),

with most sequences harboring between 12 and 21 motifs. Their distribution was uneven: motifs 7, 10, and 15 predominated and were associated with the conserved LRR consensus sequence (LxxLxLxxNxLxGxIPxxLxxLxx) (Fig. 4B), whereas motifs 12 and 14 were rare.

At the protein level, higher motif counts were generally associated with longer sequences and extensive repetition of a limited subset of motifs (Table S6). For instance, Lsat_1_v5_gn_3_60940.1 (XI-1) contained 35 motifs, representing the upper range observed, whereas Lsat_1_v5_gn_0_1380.1 (Xa) harbored only seven motifs. Despite this contrast, both proteins displayed comparable terminal motif arrangements.

Subfamily-level analyses revealed marked differences in motif abundance. Subfamilies LRR-XI-1 and -XII-1 exhibited the highest mean motif counts (>21 motifs per protein), whereas LRR-I-1, -II, -VI-2, and -Xa showed lower averages (<12 motifs per protein). Intermediate motif counts characterized larger subfamilies such as LRR-III, -V, and -VIII.

Within the kinase domain, conserved catalytic motifs corresponding to canonical subdomains were unevenly distributed among subfamilies. The GXGXXG (subdomain I) and DFG (subdomain VII) motifs were detected in 203 of the 269 LsLRR-RLKs, whereas the APE motif (subdomain VIII) was present in 173 proteins. The HRD motif (subdomain VIb) was identified in 152 members, which were accordingly classified as RD kinases; the remaining proteins were classified as non-RD kinases. Several subfamilies lacked specific catalytic motifs, including the absence of HRD in LRR-III, -V, -VI-1, and -XII-1 (Table S6). Overall, these patterns indicate a globally conserved kinase architecture accompanied by recurrent structural variation across subfamilies.

3.5. Physicochemical properties and subcellular localization prediction of the LsLRR-RLKs

The LsLRR-RLK proteins exhibited substantial biochemical variability (Fig. 5). Leucine was the most abundant amino acid residue (mean: 13.76%), followed by serine (mean: 10.42%) (Table S7). Protein length ranged from 472 to 1,640 amino acids (median: 951), with predicted molecular weights between 52.88 and 178.9 kDa (median: 104.56 kDa) and theoretical isoelectric points spanning 4.68–9.72 (median: 6.63). GRAVY values were predominantly negative (median: −0.062) (Table S8; Fig. 5A–E).

At the subfamily level, differences in protein size and biochemical properties were evident. Subfamilies LRR-XI-1, -Xb-1, and -XII-1 showed higher mean protein lengths (>1,000 aa) and molecular weights (>110 kDa), whereas LRR-II, -III, and -Xa comprised proteins with average lengths below 700 aa. Theoretical pI values also varied among subfamilies, with LRR-VII-1 presenting lower mean pI values and LRR-VI-1, -XI-2, and -XII-1 showing higher values.

GRAVY indices remained largely negative across all subfamilies, with some groups exhibiting values closer to neutrality. Subcellular localization and topology predictions indicated plasma membrane localization for most LsLRR-RLKs, whereas the AGC_MAST outgroup showed distinct biochemical properties and nuclear localization (Fig. 5F).

3.6. Gene structure, chromosomal localization, and association with SNP markers

The exon–intron organization of the 269 *LsLRR-RLK* genes and one AGC_MAST gene revealed pronounced structural heterogeneity among genes, while remaining largely conserved within subfamilies (Fig. 6). Compact gene architectures predominated in subfamilies such as LRR-III, -VII-1, -VII-2, -IX, -XI-2, and -Xb-1, which were mainly composed of one to three exons, often exceeding 2 kb in length. In contrast, subfamilies including LRR-VIII-1, -VIII-2, -V, -VI-2, and -XIIIa/XIIIb exhibited complex multi-exon structures, with up to 25–27 exons and extensive intronic regions.

Chromosomal mapping showed that *LsLRR-RLK* genes and SNP markers were unevenly distributed across the ten lettuce chromosomes (Fig. 7; Table S9). Gene density was highest on chromosomes 3 and 4, whereas SNPs were more abundant on chromosomes 5 and 2, resulting in an overall SNP-to-gene ratio of 0.71. Subfamily-level analyses revealed heterogeneous chromosomal organization patterns. The LRR-XI-1 subfamily was the most expanded and broadly distributed across nine chromosomes, with enrichment on chromosome 3, whereas LRR-XII-1 displayed a strong chromosomal bias, with approximately half of its members clustered on chromosome 2. Other subfamilies showed either wide dispersion (e.g., LRR-III) or localized accumulation, such as LRR-VIII-2 on chromosome 7 and LRR-VIII-1 on chromosome 9.

Using a ±1 Mb genomic window, 30 *LsLRR-RLK* genes (11.1% of the family) were identified in proximity to at least one SNP, yielding 57 gene–SNP associations across seven chromosomes, with an average of 1.9 SNPs per gene. These associations were not randomly distributed, with chromosome 2 showing the highest density, largely driven by clustered LRR-XII-1 loci.

3.7. Synteny analysis

In the comparison with *Arabidopsis thaliana*, a total of 115 syntenic gene pairs were identified, involving 78 *LsLRR-RLK* and 93 *Arabidopsis* genes, corresponding to 29% of the 269 genes analyzed (Fig. 8; Table S10). Among these, 30 *LsLRR-RLKs* were associated with more than one *Arabidopsis* gene, whereas 19 *Arabidopsis* genes showed syntenic relationships with multiple *LsLRR-RLKs*. Genes with the highest number of syntenic connections included *Lsat_1_v5_gn_4_163920.1* and *Lsat_1_v5_gn_6_86121.2*, each linked to four *Arabidopsis* loci. These genes were mainly located on linkage groups Lsat_1_v8_lg_4, Lsat_1_v8_lg_6, and Lsat_1_v8_lg_8, and were predominantly assigned to the LRR-III subfamily.

Comparative analysis with *Solanum lycopersicum* identified 133 syntenic gene pairs, comprising 99 *LsLRR-RLK* and 92 tomato genes, representing 36.7% of the analyzed repertoire (Fig. 8; Table S10). In this comparison, 29 *L. sativa* genes were associated with multiple tomato genes, while 34 tomato genes showed syntenic relationships with more than one *LsLRR-RLK*. The highest numbers of syntenic links were observed for *Lsat_1_v5_gn_4_89560.1*, *Lsat_1_v5_gn_7_22421.2*, and *Lsat_1_v5_gn_8_136841.1*, each connected to three tomato genes. These *LsLRR-RLKs* were distributed across several linkage groups, with recurrent representation on Lsat_1_v8_lg_4, Lsat_1_v8_lg_7, and Lsat_1_v8_lg_8, and spanned multiple subfamilies, including LRR-V, -XI-1, -III, and -VII-3.

Across both comparisons, syntenic associations involved *LsLRR-RLK* distributed throughout the *L. sativa* genome and encompassed multiple subfamilies, with several genes showing colinear relationships with more than one reference gene in each species.

3.8. Gene duplication analysis

A total of 267 duplicated *LsLRR-RLK* pairs were identified in duplication occurrences, comprising 245 whole-genome/segmental (WGD/segmental) and 22 tandem duplication events. WGD/segmental duplications represented the predominant duplication mode within the *LsLRR-RLK* family. WGD/segmental duplicates were broadly distributed across the genome, whereas tandem duplications were confined to a limited number of chromosomal regions, typically forming small local clusters (Fig. 9). All duplicated gene pairs identified were subsequently used for selective pressure analysis based on Ka/Ks estimation (Table S11).

3.9. Ka/Ks analysis and selective pressure estimation

Ka/Ks (ω) values were estimated for duplicated *LsLRR-RLK* pairs identified in the duplication analysis. Among the 267 duplicated pairs, reliable ω estimates were obtained for a subset, while 116 pairs yielded non-estimable values due to low synonymous divergence. Of

the gene pairs with reliable estimates, 143 showed $\omega < 1$, and eight exhibited $\omega > 1$ (Table S11). When stratified by duplication mode, tandem duplicates presented higher ω values (mean = 0.375; median = 0.360) than WGD/segmental duplicates (mean = 0.272; median = 0.270). All ω values remained below 1 in both categories. The lowest ω values were observed among WGD/segmental duplicates. Examples of duplicated pairs with low ω values include *Lsat_1_v5_gn_6_100781.1* and *Lsat_1_v5_gn_9_42940.1* (ω = 0.061), whereas higher ω values were observed in tandem duplicated pairs such as *Lsat_1_v5_gn_2_120461.3* and *Lsat_1_v5_gn_2_120501.1* (ω = 0.546).

3.10. Cis-regulatory elements analysis

A total of 29,876 cis-regulatory elements (CREs), representing 126 functional types grouped into eight categories, were identified within the 2 kb upstream promoter regions of *LsLRR-RLK* genes (Fig. S2; Tables S12 and S13). CRE abundance per promoter showed substantial variation across the family, with an average of 111 elements per gene.

To account for differences in effective promoter length resulting from undefined nucleotides ("N"), CRE abundance was normalized by promoter size. At the subfamily level, mean CRE density varied among *LsLRR-RLK* groups, indicating heterogeneity in promoter composition. LRR-XI-1 accumulated the highest absolute number of CREs, whereas LRR-XIII-related subfamilies and LRR-VI-2 exhibited comparatively higher CRE densities.

3.11. Transcriptomic expression profiling of *LsLRR-RLKs*

Across the analyzed transcriptomic datasets, 163 out of 269 *LsLRR-RLK* genes (60.6%) were identified as differentially expressed in at least one biological context (FDR < 0.05; $|\log_2 FC| \geq 1.5$) (Fig. 10). Among these, 113 genes were upregulated and 82 were downregulated in at least one contrast, with 32 genes displaying bidirectional regulation across different

383 conditions (Fig. 10A–C). The remaining genes did not exhibit significant transcriptional
384 changes under the evaluated conditions.

385 The number and direction of differentially expressed *LsLRR-RLK*s varied markedly
386 among biological contexts. Tissue-specific comparisons accounted for the largest fraction of
387 responsive genes, particularly root–leaf and root–stem contrasts, which were predominantly
388 characterized by transcriptional repression. In contrast, calcium deficiency showed a strong bias
389 toward transcriptional induction, with most responsive genes exhibiting positive $\log_2$FC values.
390 Biotic stress conditions also elicited substantial transcriptional responses, with partially
391 overlapping sets of *LsLRR-RLKs* detected following infection by *Rhizoctonia solani*,
392 *Sclerotinia sclerotiorum*, and *Botrytis cinerea* (Fig. 10B–C).

393 A subset of *LsLRR-RLKs* showed recurrent differential expression across multiple
394 independent contrasts. In total, 106 genes were detected as differentially expressed in two or
395 more conditions, including several genes responsive in four or more contrasts. Both stress-
396 related and developmental contexts contributed to these recurrent patterns, indicating broad
397 transcriptional responsiveness within the family.

398 Under biotic stress, infections by *S. sclerotiorum*, *R. solani*, and *B. cinerea* (48 hpi)
399 resulted in a combination of shared and pathogen-specific expression profiles, with 23 *LsLRR-*
400 *RLKs* detected in more than one fungal challenge. Abiotic stress conditions, including heat,
401 combined drought and heat, and combined drought and salinity, involved a smaller but distinct
402 subset of responsive genes, some of which displayed condition-dependent regulation across
403 multiple contrasts.

404 Responses associated with mineral nutrition varied among treatments. Calcium
405 deficiency was marked by predominant transcriptional induction, whereas phosphate-related
406 contrasts (LP vs HPC and LP vs LPC) involved fewer differentially expressed *LsLRR-RLK*s.

Developmental comparisons among roots, stems, and leaves accounted for the highest number of responsive genes and were consistently associated with extensive transcriptional repression.

4. Discussion

4.1. Identification and Classification of LRR-RLKs in *Lactuca sativa*

After stringent structural curation, 269 high-confidence LRR-RLK genes were identified in the *Lactuca sativa* genome, representing approximately 0.43% of the predicted protein-coding genes. This genomic proportion is consistent with that reported for other angiosperms and places lettuce within the expected evolutionary range for large eudicot genomes (Shiu and Bleecker 2003; Lehti-Shiu et al. 2009; Wei et al. 2015; Zhou et al. 2016; Sun et al. 2018; Meng et al. 2020; da Silva Dambroz et al. 2023). In absolute numbers, the LRR-RLK repertoire of *L. sativa* is comparable to those reported for other curated eudicot genomes, including tomato (234 genes) and common bean (230 genes), as well as woody species such as paper mulberry (236 genes) and loquat (283 genes) (Wei et al. 2015; Su et al. 2022; da Silva Dambroz et al. 2023; Yang et al. 2024). Although the relative contribution of LRR-RLKs to the lettuce proteome is lower than that observed in species such as tomato, poplar, and rice, it exceeds that reported for more ancestral plant lineages, including bryophytes and lycophytes (Liu et al. 2017), reinforcing the notion that expansion of the LRR-RLK family accompanied the diversification of land plants. Comparative analyses further indicate that variation in LRR-RLK copy number across species is primarily shaped by genome size, whole-genome duplication history, and lineage-specific retention and diversification processes (Lehti-Shiu and Shiu 2012).

Phylogenetic reconstruction based on full-length protein sequences revealed that most LsLRR-RLK subfamilies form well-supported clades, in agreement with previous LRR-RLK phylogenies reported for dicotyledonous plants (Zan et al. 2013; Wei et al. 2015; Sun et al.

2017; Liu et al. 2023). Occasional departures from strict monophyly, including isolated members within otherwise coherent subfamilies, have also been reported in *Phaseolus vulgaris* and other species and may result from annotation artifacts, domain rearrangements, or genuine evolutionary divergence (Zan et al. 2013; Magalhães et al. 2016; Cheng et al. 2021; da Silva Dambroz et al. 2023). Within this phylogenetic framework, the uneven distribution of subfamily sizes was evident, with LRR-XI-1 representing the most expanded group, whereas subfamilies such as LRR-Xb-2 and -VII-3 were sparsely represented, a pattern commonly attributed to lineage-specific duplication, retention, and loss processes (Magalhães et al. 2016; Li et al. 2018; Liu et al. 2023). It is important to note that phylogenetic relationships alone are not determinative of gene function; however, the observed concordance among tree topology, subfamily classification, and structural features supports the consistency of the inferred evolutionary framework. While this does not imply functional equivalence, it provides a rational framework for comparative functional analyses and allows the prioritization of subsets of LsLRR-RLKs as candidates for subsequent functional characterization in lettuce (Yuan et al. 2018; Su et al. 2022).

4.2. Protein architecture and biochemical features of LsLRR-RLKs

LsLRR-RLKs exhibit the canonical organization of receptor-like kinases, consisting of an extracellular domain enriched in LRR repeats with high structural plasticity and the ability to recognize a wide range of ligands, a single transmembrane helix, and a cytoplasmic serine/threonine kinase domain (Shiu and Bleecker 2001a; Liu et al. 2017). This architecture, typically preceded by a signal peptide, underlies extracellular signal perception and its subsequent intracellular transduction (Shiu and Bleecker 2001b). Across the family, the prevalence of leucine-rich regions is consistent with the structural requirements of LRR motifs, typically composed of 20–30 residues that fold into alternating β-strand and α-helical units arranged in tandem to form a characteristic solenoid-like scaffold, in which the solvent-exposed

456 β-sheet surface provides a versatile platform optimized for protein–protein and ligand
457 interactions (Zhiqi et al. 2024). Likewise, the enrichment of serine residues aligns with the
458 conserved role of serine/threonine kinases in phosphorylation-dependent signaling cascades
459 (Qin et al. 2016; Máthé et al. 2019).

460 Physicochemical profiling further supports this architecture. Predominantly negative
461 GRAVY values indicate an overall hydrophilic character, consistent with proteins possessing
462 large extracellular domains and mediating interactions at the cell surface (Kyte and Doolittle
463 1982). In parallel, subcellular localization predictions consistently supported plasma membrane
464 targeting for most LsLRR-RLKs, in agreement with the localization reported for well-
465 characterized receptor-like kinases involved in development, nutrient sensing, and immune
466 signaling in plants (Li et al. 2018; Sun et al. 2018; Cheng et al. 2021; Zhiqi et al. 2024). The
467 use of multiple prediction tools yielded concordant results and patterns comparable to those
468 described in Arabidopsis thaliana and other species (Wan et al. 2016).

469 Beyond the conserved core architecture, LsLRR-RLKs exhibited marked variability in
470 their extracellular regions. While LRR repeats predominated across all subfamilies, accessory
471 domains were detected in a subset of proteins, indicating structural diversification within the
472 family. Similar domain combinations have been reported in other plant lineages and are
473 commonly associated with lineage-specific architectural variation following gene duplication
474 (Zan et al. 2013; Yuan et al. 2018; Meng et al. 2020; Su et al. 2022).

475 Overall, the coexistence of a conserved receptor core and variable extracellular domain
476 composition highlights a structural organization compatible with functional diversification,
477 while preserving the features required for receptor-like kinase activity (Shiu et al. 2004).

478 **4.3. Structural diversification: domains, motifs, and gene architecture**

Although the canonical architecture is conserved across LsLRR-RLKs, auxiliary domain composition varies among subfamilies, consistent with domain rearrangement and duplication during receptor-like kinase evolution (Zan et al. 2013; Yuan et al. 2018; Meng et al. 2020; Su et al. 2022). Notably, malectin domains were detected in LRR-VIII-2 and malectin-like domains in LRR-I-1, both previously associated with carbohydrate binding and cell wall integrity surveillance in plants (Yang et al. 2021; Kileeg et al. 2023). Additional accessory domains, including Island, PLN00113, and PLN03150, were also identified. Such structural variation may precede processes of non-functionalization, sub-functionalization, or neo-functionalization (Gao et al. 2025).

Motif analyses indicate that LsLRR-RLKs retain a conserved catalytic and structural backbone, while exhibiting subfamily-specific variation in motif composition, distribution, and spacing, particularly within extracellular and regulatory regions. In the extracellular domain, the conserved LRR consensus motif (LxxLxLxxNxLxGxIPxxLxxLxx) underlies the formation of versatile platforms for protein–protein and ligand interactions, including pathogen-associated molecular patterns, phytohormones, and signaling peptides (Shiu and Bleecker 2001b; Liu et al. 2017). Variation in LRR copy number and the occurrence of non-LRR insertions among subfamilies are consistent with diversification in ligand recognition properties and may contribute to functional specialization by modulating ligand binding, receptor dimerization, or regulatory interactions (Nie et al. 2022; Yu et al. 2022).

Within the intracellular kinase domain, conserved catalytic motifs corresponding to canonical subdomains were widely detected (Liu et al. 2017). These include the GXGXXG and DFG motifs, which are essential for ATP binding and catalytic activity, as well as the APE motif, required for maintaining the structural integrity of the catalytic core, and the HRD motif, which defines RD kinases. The coexistence of RD kinases, commonly associated with developmental regulation, and non-RD kinases, frequently linked to pattern recognition

receptors involved in innate immunity (Dardick and Ronald 2006), together with subfamily-specific differences in motif conservation, reflects diversification within an otherwise conserved signaling framework.

Structural conservation was generally higher within subfamilies than among them, suggesting that subfamily identity constrains architectural evolution once established. In contrast, comparisons across subfamilies revealed marked differences in motif number, distribution, and combination, reinforcing the role of structural divergence in the functional radiation of LRR-RLKs. Such patterns are consistent with evolutionary models in which duplicated genes retain a conserved signaling core while gradually exploring novel regulatory or recognition capacities through structural modification (Fischer et al. 2016).

Gene structure analysis further supports the coexistence of conservation and diversification within the *LsLRR-RLK* family (Shiu and Bleecker 2003). Exon–intron organization ranged from compact, intron-poor to intron-rich architectures and was largely conserved within subfamilies (Zhang et al. 2024). This variation has been proposed to reflect distinct regulatory strategies, with intron-rich genes potentially enabling increased regulatory complexity and compact gene structures favoring rapid transcriptional responses, a pattern previously described for signaling genes involved in environmental perception (Jeffares et al. 2008; Zan et al. 2013; Zhou et al. 2016; Cheng et al. 2021). Collectively, the combined patterns of domain composition, motif organization, and gene structure indicate that *LsLRR-RLK* evolution involves the retention of a conserved signaling core alongside structural flexibility at both protein and genomic levels, supporting diversification under developmental and environmental constraints.

4.4. Genomic context: chromosomal distribution, duplication patterns, and synteny

The genomic organization of *LsLRR-RLKs* follows evolutionary patterns commonly reported for large receptor-like kinase families in plant genomes (Shiu and Bleecker 2001a). These genes are distributed across all lettuce chromosomes, with localized clustering in specific regions, a feature frequently associated with gene duplication processes in signaling gene families involved in environmental perception and defense (Zan et al. 2013; Su et al. 2022).

Duplication pattern analysis indicates that the expansion of the *LsLRR-RLK* family has been driven predominantly by large-scale events, including segmental and whole-genome duplications, which account for most duplicated gene pairs (Zan et al. 2013; Zhou et al. 2016). In contrast, tandem duplications contributed more locally, generating restricted chromosomal clusters without substantially increasing overall family size (Wei et al., 2015; Li et al., 2018). Similar distributions of duplication modes have been reported in other dicot genomes, supporting the view that large-scale duplications underpin long-term gene retention, whereas tandem duplications contribute to localized diversification within specific subfamilies (Li et al. 2018; Cheng et al. 2021; Liu et al. 2023).

Comparative synteny analyses with *Arabidopsis thaliana* and *Solanum lycopersicum* further contextualize these duplication patterns. Conserved colinear relationships involving LRR-RLK loci indicate the retention of ancestral genomic regions across eudicot lineages, while lineage-specific differences reflect chromosomal rearrangements, gene loss, and differential retention following species divergence (Coghlan et al. 2005; Wang et al. 2012). Syntenic associations were more frequently detected between lettuce and tomato than between lettuce and Arabidopsis, consistent with their closer evolutionary relationship.

Notably, several syntenic *LsLRR-RLKs* belonged to subfamilies that also exhibited local chromosomal clustering, suggesting that ancestral loci may have provided genomic contexts permissive to subsequent duplication and diversification. Comparable genomic configurations have been described for other plant gene families involved in biotic interactions, such as NB-

552 LRRs, and are generally interpreted as outcomes of lineage-specific evolutionary trajectories
553 rather than direct indicators of functional specialization (Holub 2001; Schmutz et al. 2014; Vaz
554 Bisneta and Gonçalves-Vidigal 2020).

555 Overall, the combined patterns of chromosomal distribution, duplication mode, and
556 syntenic conservation indicate that *LsLRR-RLK* evolution has proceeded through the retention
557 of ancestral genomic frameworks alongside lineage-specific expansion and rearrangement,
558 providing a genomic context for the diversification observed within this gene family in *L.*
559 *sativa*.

560 **4.5. Regulatory and transcriptional landscapes of *LsLRR-RLK*s**

561 The regulatory landscape of *LsLRR-RLK* genes indicates that promoter architecture
562 plays a central role in shaping transcriptional responsiveness across biological contexts. Cis-
563 regulatory elements (CREs) function as integration points for transcription factors acting
564 downstream of receptor-mediated signaling, and their heterogeneous distribution among
565 subfamilies reflects regulatory diversification within the *LsLRR-RLK* repertoire (Cheng et al.
566 2021; Su et al. 2022).

567 Transcriptomic analyses revealed widespread and context-dependent responsiveness of
568 *LsLRR-RLKs*, consistent with their function as upstream components of signaling networks.
569 Differential expression was detected across developmental, stress-related, and mineral nutrition
570 contrasts, with recurrent responsiveness across multiple conditions indicating that *LsLRR-RLKs*
571 constitute a flexible sensory repertoire reused in distinct physiological scenarios (Li et al. 2018;
572 Sun et al. 2018; Nie et al. 2022; Liu et al. 2023; da Silva Dambroz et al. 2023; Yang et al. 2024).
573 Tissue-specific contrasts accounted for the largest proportion of responsive genes and were
574 predominantly associated with transcriptional repression, suggesting that developmental
575 context establishes a baseline regulatory state upon which environmental and nutritional cues
576 act primarily through fine-scale modulation.

To contextualize these expression patterns, transcriptionally responsive *LsLRR-RLKs* were grouped into major biological contexts, biotic stress, abiotic stress, mineral nutrition, and development, and examined for conserved syntenic relationships with *A. thaliana*, followed by subfamily classification and phylogenetic positioning. Conserved syntenic associations involving well-characterized *AtLRR-RLKs* were prioritized to anchor functional interpretation, enabling robust block-level insights without exhaustive gene-by-gene analyses.

Across the analyzed contrasts, development-related conditions exhibited the highest number of differentially expressed *LsLRR-RLKs* and syntenic associations, followed by mineral nutrition and biotic stress, whereas abiotic stress yielded few responsive genes and no detectable syntenic pairs. Among all differentially expressed genes, *Lsat_1_v5_gn_8_46420* and *Lsat_1_v5_gn_3_53840* displayed conserved synteny with RPK1 (*AT1G69270*), a receptor implicated in the integration of developmental and stress-related cues (Hong et al. 1997; Doan et al. 2022). In parallel, *Lsat_1_v5_gn_7_115941* was syntenic to BRL1 (*AT1G55610*) and BRL3 (*AT3G13380*), receptors associated with brassinosteroid-mediated signaling.

Given that BRL1, BRL3, and their co-receptor BAK1 regulate provascular and quiescent center activity during root development (Fàbregas et al. 2013), the identification of their syntenic counterparts among biotic stress-responsive *LsLRR-RLKs* suggests that growth-related signaling pathways may also operate during immune responses, consistent with previous hypotheses linking brassinosteroid signaling to plant immunity (Kemmerling et al. 2011). Within development-related contrasts, *Lsat_1_v5_gn_2_81641* was syntenic to BAM1 (*AT5G65700*) and BAM2 (*AT3G49670*), key regulators of meristem maintenance and cell fate specification (Guo et al. 2022), supporting the conservation of core developmental signaling modules. Notably, RPK1-associated syntenic relationships extended to mineral nutrition-responsive conditions, in which the same *Lactuca* paralogous genes *Lsat_1_v5_gn_8_46420*

and *Lsat_1_v5_gn_3_53840* were differentially expressed, suggesting a broader role for RPK1-related signaling in coordinating developmental, environmental, and nutritional cues.

Overall, the integration of promoter features, transcriptional profiles, and genomic context provides a concise framework for prioritizing subsets of *LsLRR-RLK* genes for future functional studies. By emphasizing recurrent expression patterns and conserved syntenic associations across biological contexts, this approach highlights candidate receptors potentially involved in environmental adaptation and developmental regulation in lettuce.

5. Conclusion

In this study, we performed a comprehensive characterization of the LRR-RLK gene family in *Lactuca sativa*. A total of 269 *LsLRR-RLK* genes were identified and classified into 23 subfamilies, based on phylogenetic analyses consistent with the RLK-Pelle framework. All proteins exhibited the canonical LRR-RLK architecture, with conserved kinase domains and subfamily-specific variation predominantly concentrated in extracellular regions. Analyses of gene structure, conserved motifs, and domain organization revealed high uniformity within subfamilies and marked divergence among subfamilies.

Comparative synteny analyses with *Arabidopsis thaliana* and *Solanum lycopersicum* showed that a substantial fraction of *LsLRR-RLKs* is located within conserved genomic regions, suggesting the long-term retention of ancestral loci. Together with duplication analyses, these results indicate that family expansion was driven predominantly by large-scale duplication events, with a limited contribution from local duplications.

Promoter region analysis revealed extensive cis-regulatory landscapes, characterized by the presence of core regulatory elements and differential enrichment of CRE categories among subfamilies. Finally, transcriptomic analyses indicated that 163 *LsLRR-RLKs* were differentially expressed under conditions related to development, biotic and abiotic stresses,

and mineral nutrition, including genes displaying condition-specific responses as well as others with more generalist expression profiles.

Integration of expression, synteny, and phylogenetic data enabled the association of differentially expressed *Lactuca* genes with well-characterized LRR-RLKs from *Arabidopsis*, providing biologically grounded reference points for functional inference. In summary, this study establishes a curated and integrative reference for the *LsLRR-RLK* family and defines a solid framework for future functional and comparative studies in lettuce and other cultivated species.

Author contributions WAP conceptualized, conducted genomic analyses, produced the figures, supervised the project, acquired funding, and finalized the manuscript. ESO performed data collection and transcriptome analyses. KACP contributed to figure generation, particularly the phylogenetic analysis. All authors collaborated on the manuscript draft, with WAP overseeing the final revision. All authors approved the submitted version.

Acknowledgements The authors acknowledge the Department of Biology and the Graduate Program in Genetics and Plant Breeding at the Federal University of Lavras (UFLA) for providing the academic and institutional environment that supported the development of this study. The authors also acknowledge the public funding agencies that sustain the infrastructure necessary for teaching, research, and extension activities at Brazilian public universities.

Data Availability All datasets used in the differential expression analyses are publicly available. Data for *Botrytis cinerea* stress were retrieved from NCBI BioProject (PRJNA808232); mycorrhizal symbiosis from NCBI BioProject (PRJNA1043852); and *Rhizoctonia solani* stress from ArrayExpress (E-MTAB-4762). Nutrient deficiency data were obtained from GEO: GSE238167 (nitrogen/phosphate) and GSE189793 (calcium). Additional

stress-related datasets include GSE241604 (drought, heat, salinity) and GSE264560 (tissue-specific expression).

Funding statement This project did not receive specific external funding. ESO and KACP were supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) doctoral fellowships (88887.666332/2022-00 and 88887.803852/2023-00), and WAP received institutional support from FAPEMIG (APQ-00511-23).

Conflicts of interest The authors declare that there is no conflict of interest regarding the publication of this article.

Ethical approval Not applicable.

Informed consent Not applicable.

References

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed 22 March 2023].

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. 2015. The MEME Suite. Nucleic Acids Res 43(W1): W39–W49. doi:10.1093/nar/gkv416.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120. doi:10.1093/bioinformatics/btu170.

Caño-Delgado, A., Yin, Y., Yu, C., Vafeados, D., Mora-García, S., Cheng, J.-C., Nam, K.H., Li, J., and Chory, J. 2004. BRL1 and BRL3 are novel brassinosteroid receptors that function in vascular differentiation in *Arabidopsis*. Development 131(21): 5341–5351. doi:10.1242/dev.01403.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. 2020. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant 13(8): 1194–1202. doi:10.1016/j.molp.2020.06.009.

Cheng, W., Wang, Z., Xu, F., Ahmad, W., Lu, G., Su, Y., and Xu, L. 2021. Genome-wide identification of lrr-rlk family in saccharum and expression analysis in response to biotic and abiotic stress. Curr Issues Mol Biol 43(3): 1632–1651. MDPI. doi:10.3390/cimb43030116.

Coghlan, A., Eichler, E.E., Oliver, S.G., Paterson, A.H., and Stein, L. 2005. Chromosome evolution in eukaryotes: a multi-kingdom perspective. Trends in Genetics 21(12): 673–682. doi:10.1016/j.tig.2005.09.009.

Dardick, C., and Ronald, P. 2006. Plant and Animal Pathogen Recognition Receptors Signal through Non-RD Kinases. PLoS Pathog 2(1): e2. doi:10.1371/journal.ppat.0020002.

Doan, P.P.T., Kim, J.H., and Kim, J. 2022. Rapid Investigation of Functional Roles of Genes in Regulation of Leaf Senescence Using Arabidopsis Protoplasts. Front Plant Sci 13. doi:10.3389/fpls.2022.818239.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1): 15–21. doi:10.1093/bioinformatics/bts635.

Dobson, L., Reményi, I., and Tusnády, G.E. 2015. CCTOP: a Consensus Constrained TOPology prediction web server. Nucleic Acids Res 43(W1): W408–W412. doi:10.1093/nar/gkv451.

Fàbregas, N., Li, N., Boeren, S., Nash, T.E., Goshe, M.B., Clouse, S.D., de Vries, S., and Caño-Delgado, A.I. 2013. The BRASSINOSTEROID INSENSITIVE1–LIKE3 Signalosome Complex Regulates *Arabidopsis* Root Development . Plant Cell 25(9): 3377–3388. doi:10.1105/tpc.113.114462.

Fischer, I., Diévart, A., Droc, G., Dufayard, J.-F., and Chantret, N. 2016. Evolutionary Dynamics of the Leucine-Rich Repeat Receptor-Like Kinase (LRR-RLK) Subfamily in Angiosperms. Plant Physiol 170(3): 1595–1610. doi:10.1104/pp.15.01470.

Gao, Y., Qu, Q., Liu, N., Sun, M., Liu, X., Cao, Z., and Dong, J. 2025. Genome identification of the LRR-RLK gene family in maize (Zea mays) and expression analysis in response to Fusarium verticillioides infection. BMC Plant Biol 25(1): 524. doi:10.1186/s12870-025-06495-w.

Gómez-Gómez, L., and Boller, T. 2000. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Mol Cell 5(6): 1003–1011. doi:10.1016/S1097-2765(00)80265-8.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40(D1): D1178–D1186. doi:10.1093/nar/gkr944.

Gou, X., He, K., Yang, H., Yuan, T., Lin, H., Clouse, S.D., and Li, J. 2010. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in Arabidopsis thaliana. BMC Genomics 11(1): 19. doi:10.1186/1471-2164-11-19.

Guo, Z., Mei, Y., Wang, D., Xiao, D., Tang, X., Gong, Y., Xu, X., and Wang, N.N. 2022. Identification and Functional Analysis of Key Autophosphorylation Residues of Arabidopsis Senescence Associated Receptor-like Kinase. Int J Mol Sci 23(16). MDPI, ST ALBAN-ANLAGE 66, CH-4052 BASEL, SWITZERLAND. doi:10.3390/ijms23168873.

Holub, E.B. 2001. The arms race is ancient history in Arabidopsis, the wildflower. Nat Rev Genet 2(7): 516–527. doi:10.1038/35080508.

Hong, S.W., Jon, J.H., Kwak, J.M., and Nam, H.G. 1997. Identification of a Receptor-Like Protein Kinase Gene Rapidly Induced by Abscisic Acid, Dehydration, High Salt, and Cold Treatments in Arabidopsis thaliana. Plant Physiol 113(4): 1203–1212. doi:10.1104/pp.113.4.1203.

Hord, C.L.H., Chen, C., DeYoung, B.J., Clark, S.E., and Ma, H. 2006. The BAM1/BAM2 Receptor-Like Kinases Are Important Regulators of *Arabidopsis* Early Anther Development. Plant Cell 18(7): 1667–1680. doi:10.1105/tpc.105.036871.

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res 35(Web Server): W585–W587. doi:10.1093/nar/gkm259.

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. 2015. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8): 1296–1297. doi:10.1093/bioinformatics/btu817.

732 Jeffares, D.C., Penkett, C.J., and Bähler, J. 2008. Rapidly regulated genes are intron poor.
733 Trends in Genetics 24(8): 375–378. doi:10.1016/j.tig.2008.05.006.

734 Jiang, Y., Lu, X.Y., Qin, Y.L., Zhang, Y.M., and Shao, Z.Q. 2024. Genome-Wide
735 Identification and Evolutionary Analysis of Receptor-like Kinase Family Genes Provides
736 Insights into Anthracnose Resistance of Dioscorea alata. Plants 13(9). Multidisciplinary
737 Digital Publishing Institute (MDPI). doi:10.3390/plants13091274.

738 Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen,
739 J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew,
740 M., Yong, S.-Y., Lopez, R., and Hunter, S. 2014. InterProScan 5: genome-scale protein
741 function classification. Bioinformatics 30(9): 1236–1240.
742 doi:10.1093/bioinformatics/btu031.

743 Katoh, K., Rozewicki, J., and Yamada, K.D. 2018. MAFFT online service: Multiple sequence
744 alignment, interactive sequence choice and visualization. Brief Bioinform 20(4): 1160–
745 1166. Oxford University Press. doi:10.1093/bib/bbx108.

746 Kemmerling, B., Halter, T., Mazzotta, S., Mosher, S., and Nürnberger, T. 2011. A genome-
747 wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant
748 immunity. Front Plant Sci 2: 88. doi:10.3389/fpls.2011.00088.

749 Kileeg, Z., Haldar, A., Khan, H., Qamar, A., and Mott, G.A. 2023. Differential expansion and
750 retention patterns of LRR-RLK genes across plant evolution. Plant Direct 7(12).
751 doi:10.1002/pld3.556.

752 Krishnamoorthi, S., Tan, G.Z.H., Dong, Y., Leong, R., Wu, T.-Y., and Urano, D. 2024.
753 Hyperspectral imaging of liverwort Marchantia polymorpha identifies MpWRKY10 as a
754 key regulator defining Foliar pigmentation patterns. Cell Rep 43(7): 114463.
755 doi:10.1016/j.celrep.2024.114463.

756 Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.L. 2001. Predicting
757 transmembrane protein topology with a hidden markov model: application to complete
758 genomes11Edited by F. Cohen. J Mol Biol 305(3): 567–580.
759 doi:10.1006/jmbi.2000.4315.

760 Kyte, J., and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character
761 of a protein. J Mol Biol 157(1): 105–132. doi:10.1016/0022-2836(82)90515-0.

762 Lehti-Shiu, M.D., and Shiu, S.H. 2012. Diversity, classification and function of the plant
763 protein kinase superfamily. Philosophical Transactions of the Royal Society B:
764 Biological Sciences 367(1602): 2619–2639. Royal Society. doi:10.1098/rstb.2012.0003.

765 Lehti-Shiu, M.D., Zou, C., Hanada, K., and Shiu, S.-H. 2009. Evolutionary History and Stress
766 Regulation of Plant Receptor-Like Kinase/Pelle Genes . Plant Physiol 150(1): 12–26.
767 doi:10.1104/pp.108.134353.

768 Leong, R., Tan, J.J., Koh, S.S., Wu, T., Ishizaki, K., and Urano, D. 2023. G protein signaling
769 and metabolic pathways as evolutionarily conserved mechanisms to combat calcium
770 deficiency. New Phytologist 237(2): 615–630. doi:10.1111/nph.18561.

771 Lescot, M. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal
772 to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30(1): 325–327.
773 doi:10.1093/nar/30.1.325.

774 Letunic, I., and Bork, P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the
775 phylogenetic tree display and annotation tool. Nucleic Acids Res 52(W1): W78–W82.
776 doi:10.1093/nar/gkae268.

777 Letunic, I., Khedkar, S., and Bork, P. 2021. SMART: recent updates, new developments and
778 status in 2020. Nucleic Acids Res 49(D1): D458–D460. doi:10.1093/nar/gkaa937.

779 Li, X., Salman, A., Guo, C., Yu, J., Cao, S., Gao, X., Li, W., Li, H., and Guo, Y. 2018.
780 Identification and characterization of LRR-RLK family genes in potato reveal their
781 involvement in peptide signaling of cell fate decisions and biotic/abiotic stress responses.
782 Cells 7(9). MDPI. doi:10.3390/cells7090120.

783 Liu, P.L., Du, L., Huang, Y., Gao, S.M., and Yu, M. 2017. Origin and diversification of
784 leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol
785 Biol 17(1): 1–16. BioMed Central. doi:10.1186/S12862-017-0891-5.

786 Liu, S., Lei, J., Zhang, J., Liu, H., Ye, Z., Yang, J., Lu, Q., Liu, P., Chen, J., and Yang, J.
787 2023. Genome-wide identification and analysis of wheat LRR-RLK family genes
788 following Chinese wheat mosaic virus infection. Front Plant Sci 13. Frontiers Media
789 S.A. doi:10.3389/fpls.2022.1109845.

790 Magalhães, D.M., Scholte, L.L.S., Silva, N. V., Oliveira, G.C., Zipfel, C., Takita, M.A., and
791 De Souza, A.A. 2016. LRR-RLK family from two Citrus species: Genome-wide
792 identification and evolutionary aspects. BMC Genomics 17(1). BioMed Central Ltd.
793 doi:10.1186/s12864-016-2930-9.

794 Man, J., Gallagher, J.P., and Bartlett, M. 2020. Structural evolution drives diversification of
795 the large LRR-RLK gene family. New Phytologist 226(5): 1492 – 1505. Blackwell
796 Publishing Ltd. doi:10.1111/nph.16455.

797 Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire,
798 M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H.,
799 Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y.,
800 and Bryant, S.H. 2017. CDD/SPARCLE: functional classification of proteins via
801 subfamily domain architectures. Nucleic Acids Res 45(D1): D200–D203.
802 doi:10.1093/nar/gkw1129.

803 Máthé, C., Garda, T., Freytag, C., and M-Hamvas, M. 2019. The Role of Serine-Threonine
804 Protein Phosphatase PP2A in Plant Oxidative Stress Signaling—Facts and Hypotheses.
805 Int J Mol Sci 20(12): 3028. doi:10.3390/ijms20123028.

Meng, J., Yang, J., Peng, M., Liu, X., and 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 10(9): 1 – 25. MDPI AG. doi:10.3390/life10090176.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., and Bateman, A. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res 49(D1): D412–D419. doi:10.1093/nar/gkaa913.

Mu, Z., Xu, M., Manda, T., Chen, J., Yang, L., and Hwarari, D. 2024. Characterization, evolution, and abiotic stress responses of leucine-rich repeat receptor-like protein kinases (LRR-RLK) in Liriodendron chinense. BMC Genomics 25(1). BioMed Central Ltd. doi:10.1186/s12864-024-10560-3.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol 32(1): 268–274. doi:10.1093/molbev/msu300.

Nie, Z., Li, W., Deng, L., Gao, K., Liu, Q., and Zhou, Z. 2022. Comprehensive analysis of LRR-RLKs and key gene identification in Pinus massoniana resistant to pine wood nematode. Front Plant Sci 13. Frontiers Media S.A. doi:10.3389/fpls.2022.1043261.

Omasits, U., Ahrens, C.H., Müller, S., and Wollscheid, B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30(6): 884–886. doi:10.1093/bioinformatics/btt607.

Park, J.-S., Kang, M.-Y., Shim, E.-J., Oh, J., Seo, K.-I., Kim, K.S., Sim, S.-C., Chung, S.-M., Park, Y., Lee, G.P., Lee, W.-S., Kim, M., and Jung, J.-K. 2022. Genome-wide core sets of SNP markers and Fluidigm assays for rapid and effective genotypic identification of Korean cultivars of lettuce (*Lactuca sativa* L.). Hortic Res 9. doi:10.1093/hr/uhac119.

Park, S., Shi, A., Meinhardt, L.W., and Mou, B. 2023. Genome-wide characterization and evolutionary analysis of the AP2/ERF gene family in lettuce (Lactuca sativa). Sci Rep 13(1): 21990. doi:10.1038/s41598-023-49245-4.

Pink, H., Talbot, A., Carter, R., Hickman, R., Cooper, O., Law, R., Higgins, G., Yao, C., Gawthrop, F., Hand, P., Pink, D., Clarkson, J., and Denby, K. 2023, July 19. Identification of *Lactuca sativa* transcription factors impacting resistance to *Botrytis cinerea* through predictive network inference. doi:10.1101/2023.07.19.549542.

Qin, L.-X., Nie, X.-Y., Hu, R., Li, G., Xu, W.-L., and Li, X.-B. 2016. Phosphorylation of serine residue modulates cotton Di19-1 and Di19-2 activities for responding to high salinity stress and abscisic acid signaling. Sci Rep 6(1): 20371. doi:10.1038/srep20371.

Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S.M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M.A., Chovatia, M., Gepts, P., Goodstein,

D.M., Gonzales, M., Hellsten, U., Hyten, D.L., Jia, G., Kelly, J.D., Kudrna, D., Lee, R., Richard, M.M.S., Miklas, P.N., Osorno, J.M., Rodrigues, J., Thareau, V., Urrea, C.A., Wang, M., Yu, Y., Zhang, M., Wing, R.A., Cregan, P.B., Rokhsar, D.S., and Jackson, S.A. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet 46(7): 707–713. doi:10.1038/ng.3008.

Shi, Z.-X., Xiang, L., Zhao, H.-M., Yang, L.-Q., Chen, Z.-C., Pu, Y.-Q., Li, Y.-W., Luo, B., Cai, Q.-Y., Liu, B.-L., Feng, N.-X., Li, H., Li, Q.X., Tang, C., and Mo, C.-H. 2024. High-throughput single-molecule long-read RNA sequencing analysis of tissue-specific genes and isoforms in lettuce (Lactuca sativa L.). Commun Biol 7(1): 920. doi:10.1038/s42003-024-06598-4.

Shiu, S.-H., and Bleecker, A.B. 2001a. Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. Proceedings of the National Academy of Sciences 98(19): 10763–10768. doi:10.1073/pnas.181141598.

Shiu, S.-H., and Bleecker, A.B. 2001b. Plant Receptor-Like Kinase Gene Family: Diversity, Function, and Signaling. Science's STKE 2001(113). doi:10.1126/stke.2001.113.re22.

Shiu, S.H., and Bleecker, A.B. 2003. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol 132(2): 530–543. American Society of Plant Biologists. doi:10.1104/pp.103.021964.

Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F.X., and Li, W.H. 2004. Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell 16(5): 1220–1234. American Society of Plant Biologists. doi:10.1105/tpc.020834.

da Silva Dambroz, C.M., Aono, A.H., de Andrade Silva, E.M., and Pereira, W.A. 2023. Genome-wide analysis and characterization of the LRR-RLK gene family provides insights into anthracnose resistance in common bean. Sci Rep 13(1). Nature Research. doi:10.1038/s41598-023-40054-3.

Stothard, P. 2000. The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. Biotechniques 28(6): 1102–1104. doi:10.2144/00286ir01.

Su, Y., Peng, X., and Shen, S. 2022. Identification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in paper mulberry and their potential roles in response to cold stress. Comput Biol Chem 97. Elsevier Ltd. doi:10.1016/j.compbiolchem.2022.107622.

Sun, J., Li, L., Wang, P., Zhang, S., and 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(1). BioMed Central Ltd. doi:10.1186/s12864-017-4155-y.

Sun, R., Wang, S., Ma, D., and Liu, C. 2018. Genome-Wide Analysis of LRR-RLK Gene Family in Four Gossypium Species and Expression Analysis during Cotton Development and Stress Responses. Genes (Basel) 9(12): 592. doi:10.3390/genes9120592.

Sun, Y., Li, L., Macho, A.P., Han, Z., Hu, Z., Zipfel, C., Zhou, J.-M., and Chai, J. 2013. Structural Basis for flg22-Induced Activation of the *Arabidopsis* FLS2-BAK1 Immune Complex. Science (1979) 342(6158): 624–628. doi:10.1126/science.1243825.

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40(7): 1023–1025. doi:10.1038/s41587-021-01156-3.

Tsirigos, K.D., Peters, C., Shu, N., Käll, L., and Elofsson, A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res 43(W1): W401–W407. doi:10.1093/nar/gkv485.

Vangelisti, A., Cela, F., Simoni, S., Avio, L., Turrini, A., Sbrana, C., Incrocci, L., Pardossi, A., Rogo, U., Natali, L., Cavallini, A., and Giordani, T. 2024. Arbuscular mycorrhizal symbiosis allows to overcome phosphorus deficiency in soilless culture of Lactuca sativa: A phenotypic and transcriptomic analysis. Sci Hortic 334: 113324. doi:10.1016/j.scienta.2024.113324.

Vaz Bisneta, M., and Gonçalves-Vidigal, M.C. 2020. Integration of anthracnose resistance loci and RLK and NBS-LRR-encoding genes in the *Phaseolus vulgaris* L. genome. Crop Sci 60(6): 2901–2918. doi:10.1002/csc2.20288.

Verwaaijen, B., Wibberg, D., Winkler, A., Zrenner, R., Bednarz, H., Niehaus, K., Grosch, R., Pühler, A., and Schlüter, A. 2019. A comprehensive analysis of the Lactuca sativa, L. transcriptome during different stages of the compatible interaction with Rhizoctonia solani. Sci Rep 9(1): 7221. doi:10.1038/s41598-019-43706-5.

Wan, S., Mak, M.-W., and Kung, S.-Y. 2016. Mem-mEN: Predicting Multi-Functional Types of Membrane Proteins by Interpretable Elastic Nets. IEEE/ACM Trans Comput Biol Bioinform 13(4): 706–718. doi:10.1109/TCBB.2015.2474407.

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., and Paterson, A.H. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40(7): e49–e49. doi:10.1093/nar/gkr1293.

Wei, Z., Wang, J., Yang, S., and Song, Y. 2015. Identification and expression analysis of the *LRR-RLK* gene family in tomato (*Solanum lycopersicum*) Heinz 1706. Genome 58(4): 121–134. doi:10.1139/gen-2015-0035.

Wolfe, D., Dudek, S., Ritchie, M.D., and Pendergrass, S.A. 2013. Visualizing genomic information across chromosomes with PhenoGram. BioData Min 6(1). doi:10.1186/1756-0381-6-18.

Wu, Y., Xun, Q., Guo, Y., Zhang, J., Cheng, K., Shi, T., He, K., Hou, S., Gou, X., and Li, J. 2016. Genome-Wide Expression Pattern Analyses of the Arabidopsis Leucine-Rich

Repeat Receptor-Like Kinases. Mol Plant 9(2): 289–300. Cell Press. doi:10.1016/j.molp.2015.12.011.

Yang, H., Wang, D., Guo, L., Pan, H., Yvon, R., Garman, S., Wu, H.-M., and Cheung, A.Y. 2021. Malectin/Malectin-like domain-containing proteins: A repertoire of cell surface molecules with broad functional potential. The Cell Surface 7: 100056. doi:10.1016/j.tcsw.2021.100056.

Yang, M., Min, T., Manda, T., Yang, L., and Hwarari, D. 2024. Genomic Survey of LRR-RLK Genes in Eriobotrya japonica and Their Expression Patterns Responding to Environmental Stresses. Plants 13(17). Multidisciplinary Digital Publishing Institute (MDPI). doi:10.3390/plants13172387.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol 24(8): 1586–1591. doi:10.1093/molbev/msm088.

Yu, C.-S., Chen, Y.-C., Lu, C.-H., and Hwang, J.-K. 2006. Prediction of protein subcellular localization. Proteins: Structure, Function, and Bioinformatics 64(3): 643–651. doi:10.1002/prot.21018.

Yu, J., Zhang, B., Liu, S., Guo, W., Gao, Y., and Sun, H. 2022. Genome-wide characterisation, evolution and expression analysis of the leucine-rich repeat receptor-like kinase (LRR-RLK) gene family in cucumbers. Plant Protection Science 58(2): 125 – 138. Czech Academy of Agricultural Sciences. doi:10.17221/131/2021-PPS.

Yu, Y., Song, W., Zhai, N., Zhang, S., Wang, J., Wang, S., Liu, W., Huang, C.-H., Ma, H., Chai, J., and Chang, F. 2023. PXL1 and SERKs act as receptor–coreceptor complexes for the CLE19 peptide to regulate pollen development. Nat Commun 14(1): 3307. doi:10.1038/s41467-023-39074-4.

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

Zan, Y., Ji, Y., Zhang, Y., Yang, S., Song, Y., and Wang, J. 2013. Genome-wide identification, characterization and expression analysis of populus leucine-rich repeat receptor-like protein kinase genes. BMC Genomics 14(1). doi:10.1186/1471-2164-14-318.

Zhang, Z., Ye, F., Hu, K., Luo, T., and Miao, Z. 2024. New insights into evolution and functional diversification of Camellia sinensis LRR-RLKs. Physiology and Molecular Biology of Plants 30(5): 851 – 866. Springer. doi:10.1007/s12298-024-01458-1.

Zhiqi, H., Tingyi, W., Dongdong, C., Lan, S., Guangheng, Z., Qian, Q., and Li, Z. 2024. Leucine-Rich Repeat Protein Family Regulates Stress Tolerance and Development in Plants. Rice Sci. doi:10.1016/j.rsci.2024.12.003.

956 Zhou, F., Guo, Y., and Qiu, L.J. 2016. Genome-wide identification and evolutionary analysis
957 of leucine-rich repeat receptor-like protein kinase genes in soybean. BMC Plant Biol
958 16(1). BioMed Central Ltd. doi:10.1186/s12870-016-0744-1.

959 Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T., and Felix, G.
960 2006. Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts
961 Agrobacterium-Mediated Transformation. Cell 125(4): 749–760.
962 doi:10.1016/j.cell.2006.03.037.

963

964 **Figure Captions**

965 **Figure 1.** Integrated analytical workflow for LRR-RLK analysis in *Lactuca sativa.* Integrated
966 workflow employed for the genome-wide identification, classification, and multi-level
967 characterization of the *LRR-RLK* gene family in *Lactuca sativa*. The pipeline summarizes
968 successive analytical steps, including sequence retrieval, phylogenetic classification, domain
969 and motif architecture analysis, physicochemical characterization, gene structure and
970 chromosomal mapping, cis-regulatory element identification, synteny and duplication analyses,
971 selective pressure inference, and transcriptomic profiling under stress and developmental
972 conditions. The main bioinformatic tools and databases used at each step are indicated in
973 parentheses.

974 **Figure 2.** Phylogenetic relationships of LRR-RLK proteins in *Lactuca sativa*. Phylogenetic
975 analyses were performed using full-length amino acid sequences of curated LsLRR-RLK
976 proteins. Sequences were aligned with MAFFT v7 using the L-INS-i strategy, and maximum-
977 likelihood trees were inferred with IQ-TREE v2, with substitution models selected
978 automatically by ModelFinder and branch support assessed using 1,000 ultrafast bootstrap
979 replicates. The tree shown corresponds to the *L. sativa*–specific phylogenetic reconstruction.
980 The AGC_MAST serine/threonine kinase Lsat_1_v5_gn_0_34000.1 was used as an outgroup.
981 LRR-RLK subfamilies are indicated by colored branches according to the classification
982 framework proposed by Lehti-Shiu and Shiu (2012). Bootstrap support values are shown for
983 major nodes.

984 **Figure 3.** Domain architecture and phylogenetic context of LsLRR-RLK proteins. (A)
985 Maximum-likelihood phylogenetic tree of LsLRR-RLK proteins inferred from full-length
986 amino acid sequences, illustrating subfamily-level relationships. (B) Linear representation of
987 protein domain architectures aligned according to the phylogenetic order shown in panel A,
988 highlighting the organization and distribution of conserved domains along each LsLRR-RLK

sequence. Domain annotations were obtained using Pfam and the NCBI Conserved Domain Database (CDD), and the integrated visualization of phylogeny and domain architecture was generated using TBtools. Conserved domains and structural features are indicated by colored boxes, as shown in the legend.

Figure 4. Conserved motif organization of LsLRR-RLK proteins. (A) Phylogenetic tree of LsLRR-RLK proteins, showing the relative position of each sequence; (B) Linear representation of LsLRR-RLK protein architectures, illustrating the distribution and arrangement of the 15 conserved motifs identified by MEME along each protein sequence; (C) Sequence logos of the 15 conserved motifs, highlighting amino acid conservation patterns. The figure was generated using motif output files from the MEME Suite and visualized with TBtools.

Figure 5. Physicochemical properties and subcellular localization of LsLRR-RLK proteins. (A) Percentage composition of amino acids for each LsLRR-RLK protein. (B) Physicochemical properties, including protein length (amino acids), molecular weight (kDa), theoretical isoelectric point (pI), and grand average of hydropathicity (GRAVY). (C) Prediction of transmembrane helices based on four independent tools (TMHMM, TOPCONS, Protter, and CCTOP). (D) Prediction of signal peptide presence using TMHMM, TOPCONS, Protter, and SignalP 6.0. (E) Predicted subcellular localization of LsLRR-RLK proteins. Proteins are ordered according to their position in the phylogenetic tree, allowing direct comparison between phylogenetic relationships and physicochemical and localization features.

Figure 6. Gene structure organization of *LsLRR-RLK* genes in *Lactuca sativa*. The exon–intron architecture of each *LsLRR-RLK* gene is shown schematically, with exons represented by red boxes and introns indicated by black lines. Predicted 5′ and 3′ untranslated regions (UTRs) are shown in blue. Genes are ordered according to their position in the phylogenetic tree, allowing direct comparison between gene structure and phylogenetic relationships.

Figure 7. Chromosomal distribution of *LsLRR-RLK* genes in *Lactuca sativa*. The physical positions of 269 *LsLRR-RLK* genes are shown across the nine chromosomes, with colors indicating subfamily classification. Previously reported SNP markers mapped to the *L. sativa* genome (Park et al. 2022) are indicated by triangles. Six genes, including the AGC_MAST outgroup, were located on unanchored scaffolds and are shown on Chr0. The figure was generated using the Phenogram tool.

Figure 8. Syntenic relationships of *LRR-RLK* genes among *Arabidopsis thaliana*, *Lactuca sativa*, and *Solanum lycopersicum*. Chromosomes are represented as horizontal bars for each species: *A. thaliana* (top), *L. sativa* (middle), and *S. lycopersicum* (bottom). Lines indicate syntenic links involving *LRR-RLK* genes identified among the three genomes.

Figure 9. Segmental and tandem duplication of *LsLRR-RLK* genes. A Circos plot illustrates duplicated gene pairs identified across the genome, with chromosomes represented as gray segments. Blue links indicate segmental duplication events, whereas red links represent tandem duplication events between *LsLRR-RLK* genes.

Figure 10. Expression profiles of *LsLRR-RLK* genes under stress and developmental conditions. (A) Circular heatmap showing $\log_2$ fold-change ($\log_2$FC) values of differentially expressed *LsLRR-RLK* genes across multiple contrasts, including tissue comparisons (stem vs leaf, root vs stem, root vs leaf), nutrient availability (N and P), abiotic stresses (calcium deficiency, heat, drought, salt, and combined stresses), and biotic interactions with *Sclerotinia sclerotiorum*, *Rhizoctonia solani*, and *Botrytis cinerea*. (B) Bar plot summarizing the number of upregulated (UP) and downregulated (DOWN) *LsLRR-RLK* genes detected in each contrast. (C) UpSet plot illustrating the overlap of differentially expressed *LsLRR-RLK* genes among contrasts, highlighting shared and condition-specific transcriptional responses. Color scales represent $\log_2$FC values.