



**HAL**  
open science

# Analyses of Lysin-motif Receptor-like Kinase (LysM-RLK) Gene Family in Allotetraploid *Brassica napus* L. and Its Progenitor Species: An In Silico Study

Amin Abedi, Zahra Hajiahmadi, Mojtaba Kordrostami, Qassim Esmaeel, Cédric Jacquard

## ► To cite this version:

Amin Abedi, Zahra Hajiahmadi, Mojtaba Kordrostami, Qassim Esmaeel, Cédric Jacquard. Analyses of Lysin-motif Receptor-like Kinase (LysM-RLK) Gene Family in Allotetraploid *Brassica napus* L. and Its Progenitor Species: An In Silico Study. *Cells*, 2022, 11 (1), pp.37. 10.3390/cells11010037. hal-03511666

HAL Id: hal-03511666

<https://hal.univ-reims.fr/hal-03511666>

Submitted on 5 Jan 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License

Article

# Analyses of Lysin-motif Receptor-like Kinase (*LysM-RLK*) Gene Family in Allotetraploid *Brassica napus* L. and Its Progenitor Species: An In Silico Study

Amin Abedi <sup>1,†</sup> , Zahra Hajiahmadi <sup>1,†</sup>, Mojtaba Kordrostami <sup>2,\*</sup>, Qassim Esmaeel <sup>3</sup>  and Cédric Jacquard <sup>3,\*</sup> 

<sup>1</sup> Department of Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht 41635-1314, Iran; abedi.amin@yahoo.com (A.A.); z.hajiahmadi1366@gmail.com (Z.H.)

<sup>2</sup> Nuclear Agriculture Research School, Nuclear Science and Technology Research Institute (NSTRI), Atomic Energy Organization of Iran (AEOI), Karaj 31485-498, Iran

<sup>3</sup> Université de Reims Champagne-Ardenne, RIBP EA4707 USC INRAE 1488, SFR Condorcet FR CNRS 3417, 51100 Reims, France; qassim.esmaeel@univ-reims.fr

\* Correspondence: mkordrostami@aeoi.org.ir (M.K.); cedric.jacquard@univ-reims.fr (C.J.); Tel.: +98-9177015872 (M.K.); +33-3-26-91-34-36 (C.J.)

† These authors contributed equally to this work.

**Abstract:** The LysM receptor-like kinases (LysM-RLKs) play a crucial role in plant symbiosis and response to environmental stresses. *Brassica napus*, *B. rapa*, and *B. oleracea* are utilized as valuable vegetables. Different biotic and abiotic stressors affect these crops, resulting in yield losses. Therefore, genome-wide analysis of the LysM-RLK gene family was conducted. From the genome of the examined species, 33 LysM-RLK have been found. The conserved domains of *Brassica* LysM-RLKs were divided into three groups: LYK, LYP, and LysMn. In the *Brassica LysM-RLK* gene family, only segmental duplication has occurred. The Ka/Ks ratio for the duplicated pair of genes was less than one indicating that the genes' function had not changed over time. The *Brassica LysM-RLKs* contain 70 cis-elements, indicating that they are involved in stress response. 39 miRNA molecules were responsible for the post-transcriptional regulation of 12 *Brassica LysM-RLKs*. A total of 22 SSR loci were discovered in 16 *Brassica LysM-RLKs*. According to RNA-seq data, the highest expression in response to biotic stresses was related to BnLYP6. According to the docking simulations, several residues in the active sites of BnLYP6 are in direct contact with the docked chitin and could be useful in future studies to develop pathogen-resistant *B. napus*. This research reveals comprehensive information that could lead to the identification of potential genes for *Brassica* species genetic manipulation.

**Keywords:** bioinformatics; codon usage bias; expression pattern; in silico study; modeling; molecular docking



**Citation:** Abedi, A.; Hajiahmadi, Z.; Kordrostami, M.; Esmaeel, Q.; Jacquard, C. Analyses of Lysin-motif Receptor-like Kinase (*LysM-RLK*) Gene Family in Allotetraploid *Brassica napus* L. and Its Progenitor Species: An In Silico Study. *Cells* **2022**, *11*, 37. <https://doi.org/10.3390/cells11010037>

Academic Editors: Francesco Mercati and Suleyman Allakhverdiev

Received: 2 November 2021

Accepted: 20 December 2021

Published: 23 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Because plants are stationary, they are subjected to a variety of biotic and abiotic stresses throughout their lives. Plants have been developed their immune strategies using signal transduction from the site of infection [1]. Immune receptors are used by plants to detect and respond to invading pathogens [2]. Plants' immune receptors are classified as either nucleotide-binding leucine-rich repeat receptors (NLR) or pattern recognition receptors (PRRs) [3]. Accordingly, NLR and PRR receptors are known as resistance gene analogs (RGAs). NLRs play a major role in plant disease resistance and are also known as resistance genes (R genes) [4].

PRRs are the main line of defense against infections. They are located on the cell membrane and belong to a receptor kinase family. They contain an intracellular kinase domain, a transmembrane domain, as well as an extracellular ligand-binding domain [5]. The extracellular domain recognizes molecular-associated molecular patterns (MAMPs). MAMPs are cell envelope components such as lipopolysaccharide (LPS), flagellin, chitin,

$\beta$ -glucans, peptidoglycan, and ergosterols [2,6]. Proteins containing Lysin-motif (LysM) are PRRs that recognize MAMPs [3,7]. Plant PRRs are classified into two groups: receptor-like proteins (RLPs) and receptor-like kinases (RLKs). RLKs contain an intracellular kinase domain, a transmembrane domain, and also an extracellular ligand-binding domain that is involved in signal transduction while RLPs lack intracellular regions [8]. Plant RLKs have a similar structure to animal receptor tyrosine kinases. Several extracellular domains have been discovered in plants, but none have been identified in animals [9]. In their extracellular regions, plant RLKs exhibit a diverse set of domains [10]. Plant RLKs can be divided into 14 groups, which include wall-associated kinase (WAK), receptor-like kinase in flowers (RKF), CRINKLY-like (CR-like), *Catharanthus roseus* like (CrRLK), the domain of unknown function 26 (DUF26), lectin (C-Lectin and L-Lectin), leucine-rich repeat (LRR), lysin motif (LysM), an extension like proline-rich extensin like (PERK), leaf rust kinase-like (LRK), thaumatin, self-incompatibility domain (S-domain), and unknown receptor kinase (URK) with various functions [9]. Few RLKs are known to have crucial roles in plant defense such as lysin motif- receptor-like kinases (LysM-RLKs), one of the most important groups implicated in plant defense response. LysM has a highly conserved  $\beta\alpha\alpha\beta$  secondary structure containing about 50 amino acids that bind to chitin and peptidoglycans [11]. It was first discovered in lysozyme [12] of a bacteriophage and later shown to be present in a wide range of eukaryotes and prokaryotes [13,14]. LysM family contains LysM-containing receptor-like kinase (LYK) and LysM-containing receptor-like proteins (LYP) which are widely distributed in the plant kingdom [15,16]. LYKs are important in plant-pathogen interactions because they activate the immune system by sensing pathogen entrance into the host cell [17]. The LysM motif is found in the extracellular region of LYKs [18]. LYKs are essential mediators of innate immunity against pathogens in a broad range of plant species. Investigation of the properties of the plant LysM showed that plant genomes have at least 11 distinct types of this motif, which are highly diverse in plants and at least six types of LysM motif in LysM kinase proteins and five other types in non-kinase LysM proteins have been identified [19]. The first receptor protein (*Oryza sativa* chitin elicitor-binding protein-*OsCEBiP*) was identified in rice which contains two extracellular LysM domains and a transmembrane LysM domain. Silencing of *OsCEBiP* prevented the plant from responding to chitin [15]. Wheat and barley have orthologs of *OsCEBiP* that are engaged in the plant defence system. The *Mycosphaerella graminicola* pathogen caused disease symptoms in wheat lines that were knocked down for *TaCEBiP* (*Triticum aestivum* chitin elicitor binding protein) using the virus-induced gene silencing (VIGS) approach [20]. In the barley lines knocked down for *HvCEBiP* (*Hordeum vulgare* chitin elicitor binding protein) via the VIGS, increased lesions owing to *Magnaporthe oryzae* infection were also observed [21]. Similarly, *OsCERK1* (*Oryza sativa* chitin elicitor receptor kinase 1) is a LysM receptor-like protein kinase with three lysine motifs and a kinase domain that is required for signal transduction in rice. However, *OsCEBiP* is the main chitin-binding protein that uses *OsCERK1* to activate the chitin-stimulated immune response [22–24]. *OsCERK1* transmits the chitin signal to the cell, which has been sensed by *OsCEBiP*. *CERK1* directly binds to chitin by three extracellular LysM domains without any other proteins indicating that it is the main chitin-binding protein in plants [25]. The activation of Mitogen-activated protein kinase (MAPK), the formation of reactive oxygen species (ROS), and the expression of defense genes are all part of the immunological response caused by fungal chitin [26]. According to certain studies, *AtCERK* (*Arabidopsis thaliana* chitin elicitor receptor kinase) has a dual role in biotic and abiotic stress signaling [27]. In *Arabidopsis*, *Atlyk1* and *Atlyk4* bind to chitin and transmit extracellular signals to the cell, activating downstream pathways of disease resistance [28]. *AtLYM1* and *AtLYM3* are similar to *OsCEBiP* and have a role in binding to peptidoglycans [29]. *AtLYM2* has been reported to increase plant resistance against *Botrytis cinerea* and *Alternaria brassicicola* [30,31]. *VvLYK1-1* and *VvLYK1-2* in grapevine are involved in immunity induced by chitin and chitosan, and hence *VvLYK1-1* is implicated in *Erysiphe necator* resistance [32].

In addition to the plant immune response, LysM-RLK genes are also involved in the plant and arbuscular mycorrhizal (AM) fungi interaction. Studies on chickpeas have shown that PsLYK9 is directly involved in the perception of long- and short-chain oligosaccharides in the hydrolyzed cell walls of the fungus and plays an important role in the immune response of chickpeas to fungal pathogens. PsLYK9, on the other hand, is involved in the symbiosis development of AM fungi, so that silencing *PsLYK9* reduces levels of colonization by mycorrhizal fungi [33]. In tomatoes, SILYK1 and SILYK13 are involved in the chitin-induced immune response and cell death, respectively. While SILYK10 and SILYK13 participate in the regulation of AM symbiosis [34,35].

*Brassica napus* L. ( $2n = 4x = 38$ ) is one of the most important allopolyploid oilseed crops derived from the hybridization of *B. oleracea* ( $2n = 2x = 18$ ), and *B. rapa* ( $2n = 2x = 20$ ). These species are used as important vegetables in the form of sauces, oil, and fodder, among many other items. Their crops are subjected to biotic and abiotic stressors throughout their life cycle, resulting in yield reductions. *Alternaria* spp., *Fusarium oxysporum*, *Albugo candida*, and *Leptosphaeria maculans* are the most prevalent fungal diseases that affect *Brassica* [36]. *Brassica* crops are susceptible to bacterial rot disease as well [37]. Similarly, salinity and drought among abiotic stresses have the highest impact on *Brassica* yield [38]. To boost crop productivity, stress-tolerant *Brassica* can be developed through genetic engineering. In Arabidopsis, rice, lotus, sweet orange, potato, wheat, Chinese white pear, as well as other crops, the Lysin-motif Receptor-Like Kinase (LysM-RLK) family has been explored [8,17,39–44]. However, no genome-wide association studies in *B. napus* have been conducted to date. As a result, functional investigations of *LysM-RLKs* in *Brassica* have been investigated using bioinformatics techniques, given the importance of *Brassica* and the decrease of its products owing to biotic and abiotic stressors.

## 2. Materials and Methods

### 2.1. In Silico Identification of LysM-RLK Genes

The HMM profile of the LysM domain (Pfam ID: PF01476) was retrieved from the Pfam database [45], and the HmmerSearch tool [46] was used to determine *Brassica* LysM-RLK proteins in the Ensembl Plants database to determine the LysM-RLK gene family in *B. napus*, *B. oleracea*, and *B. rapa*. Three domains including F-box-like (PF12937), protein tyrosine kinase (PF07714), and protein kinase domain (PF00069) were identified. The default parameters include significance E-values of 0.01 for sequence and 0.03 for hit matches, as well as reporting E-values of 1 for both sequences and hit. The ProtParam tool of the ExPASy bioinformatics resource portal [47] was used to compute the molecular weight, length, and theoretical isoelectric points of *Brassica* LysM-RLK. CELLO and DeepLoc were used to predict protein cellular localization [48,49].

### 2.2. Phylogenetic Relationships of Brassica LysM-RLK Gene Family

To study the evolutionary relationships of the *LysM-RLK* gene family, full-length protein sequence alignments of *B. napus* (Bn), *B. oleracea* (Bo), *B. rapa* (Br), *Arabidopsis thaliana* (At), *Brachypodium distachyon* (Bd), *Oryza sativa* (Os), and *Vitis vinifera* (Vv) were performed using ClustalX 2.0.8 software. The Neighbor-joining (NJ) method with 1000 bootstraps was used to generate a phylogenetic tree of LysM-RLK proteins using MEGA 7 [50] and the p-distance model [51].

### 2.3. Investigation of Chromosome Localization, Gene Duplication, and Selection Pressure of LysM-RLK Members

The coding sequence (CDS) of the examined genes was retrieved from the Ensembl Plants database using the biomart program [52] to investigate duplication and selection pressure. Tandem duplication is defined as the duplication of genes on the same chromosome separated by no more than 10 genes [53]. Similarly, two criteria were utilized to detect segmental duplication: the aligned region's identity had to be larger than 90% and the alignment coverage had to be greater than 90% [54]. Synonymous (Ks) and non-

synonymous (Ka) substitution rates were evaluated using DnaSP ver. 5 software [55] to establish the kind of selection pressure. TBtools [56] was used to determine the location of genes on chromosomes and the duplication relationship among them.

#### 2.4. Exon-Intron Structure and Conserved Motifs of *BLysM-RLK*

The Multiple Em for Motif Elicitation (MEME 5.0.5) algorithm was used to find specific *LysM-RLK* gene motifs [57]. Twenty motifs with a minimum and maximum length of motifs 6 and two hundred amino acids have been considered. These findings were shown using the TBtools software [56]. The GFF3 file linked to the three *Brassica* species was retrieved from the Ensemble Plants database and the appropriate analyses were done using TBtools software [56] to illustrate the exon-intron structure of the examined genes.

#### 2.5. The Prediction of Cis-Regulatory Elements, Simple Sequence Repeats (SSR) Markers, and *BLysM-RLK*-Targeted miRNAs

PlantCare [58] was used to identify cis-acting regulatory elements of 1500 bp upstream of the initiation codon (ATG) of the *LysM-RLK* genes from the Ensemble Plants database [52]. SSR markers in *BLysM-RLK* genes were discovered using the BatchPrimer3v1.0 server [59]. CD sequences of them were evaluated in the psRNATarget database using default parameters to detect in *BLysM-RLK*-targeted miRNAs.

#### 2.6. Codon Usage Bias Analysis

CodonW 1.4.2 was used to analyze the sequences for frequency of optimal codons (FOP), codon adaptation index (CAI), GC content, effective codon number (ENC), GC content at the third site position of a codon (GC3s), and relative synonymous codon usage (RSCU) for *Brassica LysM-RLK* [60]. The statistical analysis was carried out using Excel software.

#### 2.7. RNA-Seq Analysis of *Brassica LysM-RLK* Genes

The transcript data for flower, leaf, root, silique, and stem tissues as well as dehydration stress at 1 and 8 h after treatment and ABA (25 M), cold (4 °C), and salinity (200 mM), stresses at 4 and 24 h after treatment were related to the study of Zhang et al. [61] with the project ID CRA001775 [62]. FastQC software [48] was used for the initial quality analysis on FastQ files, and then the raw sequence data was preprocessed and adapter sequences, low-quality reads, and duplicate mapping reads were filtered using Trimmomatic on Linux [49]. The preprocessed FastQ files were aligned to the *Brassica napus* reference genome using STAR [50]. The counts obtained from STAR normalized to transcript per million (TPM). Log<sub>2</sub> (TPM + 1) used to generate the heatmap utilizing TBtools [63]. Clustering the data was performed using the Pearson correlation coincident and the complete linkage method. Similarly, the BrassicaEDB database was used to study the expression of *BnLysM-RLK* genes in response to fungal infections such as *Leptosphaeria maculans* and *Sclerotinia sclerotiorum*. Expression data related to the *Leptosphaeria Maculans* inoculation is available in the NCBI with the project ID number PRJNA311316. In total, they sequenced 36 samples (18 from the resistant (LepR1) genotype and 18 from the susceptible genotype (Westar)). Samples were collected 0, 3, 7, and 11 days post-inoculation in triplicate. RNA-seq data with accession number PRJNA274853 publicly available on the NCBI SRA database were mined and analyzed for expression patterns of the rapeseed *LysM-RLK* genes in response to *S. sclerotiorum* infection. The experiment consisted of 24 samples containing susceptible (J902) and resistant (J964) genotypes and was sampled at 24, 48, and 96 h after treatment with three biological replications.

#### 2.8. Structural Modeling and Validation

Iterative template-based fragment assembly simulations were used to create the full-length atomic structures of BnLYP6 proteins to forecast protein structures on the I-TASSER server [64]. The top models from I-TASSER were refined using the ModRefinder soft-

ware [65]. Ramachandran plot has been applied to confirm the predicted structures by measuring the backbone dihedral phi ( $\phi$ ) and psi ( $\Psi$ ) angles with the PROCHECK module of the PDBSum server [66].

### 2.9. Molecular Docking

The chitin ligand structure was retrieved from the PubChem database [67] and converted to PDB format using Discovery Studio software. An improved version of the COACH server (COACH-D) was utilized to discover protein-ligand interaction sites [68]. To suggest protein-ligand binding sites, the aforementioned server employs five approaches, four of which are template-based, including TM-SITE [69], COFACTOR [70], and FIND-SITE [71] while the last method (ConCavity) is based on structure [72]. The results of each approach were then combined using the COACH algorithm [69]. The ligand-enzyme interaction was studied using AutoDock v4.2.6 [73]. The Auto Grid application, which was created with AutoDock, was used to create grid maps. The grid box sizes for x, y, and z were set to 82, 90, and 120, respectively. The grid centers for x, y, and z were set at 73.866, 76.789, and 68.402, respectively, with a grid spacing of 0.375. To find the best conformers, the Lamarckian Genetic Algorithm (LGA) was used. During the docking process, a limit of 200 conformers was considered for the ligand. The default AutoDock4 parameters were used for the majority of docking processes [73]. The maximum number of tests was set at 2,500,000, the population size was set at 150, the maximum number of generations was set at 27,000, the maximum number of automatically surviving top individuals was set at 1, the gene mutation rate was set at 0.02 and the crossover rate was set at 0.8. The interaction of enzymes and substrates has been illustrated in 2D and 3D using Discovery Studio Visualizer and Chimera softwares [74].

## 3. Results

### 3.1. Identification of *Brassica* LysM-RLK Genes

In the current investigation, 33 *LysM-RLK* genes were discovered (17 in *B. napus*, 8 in each of *B. rapa* and *B. oleracea*). The prefix Bn, Bo, and Br, as well as the protected domain discovered in each gene, were used to label the identified *LysM-RLK* genes. The chromosomal location of the genes has been used to estimate the gene number. They were divided into three groups based on their specific domains including LYK (5 in *B. napus*, 2 in each of *B. rapa* and *B. oleracea*), LYP (10 in *B. napus*, 5 in each of *B. rapa* and *B. oleracea*), and LysMn (2 in *B. napus*, 1 in each of *B. rapa* and *B. oleracea*) (Table 1). LYKs are made up of LysM and protein kinase domains, according to supplementary file 1 (Table S1). LysM domains have been discovered in LYPs. Some LYPs are transmembrane, whereas others use a glycosylphosphatidylinositol anchor to bind to the membrane. LysMn proteins have an F box-like domain with extracellular or plasma membrane localization. The physicochemical characteristics of LysM-RLK were investigated using the ProtParam tool. The length of these 33 BnATGs protein sequences ranged from 260 amino acids (BnLysMn1-2, BrLysMn, and BoLysMn) to 665 amino acids (BnLysMn1-2, BrLysMn, and BoLysMn) (BnLYK3 and BoLYK1). The molecular weights of LysM-RLK proteins ranged from 4.08 to 72.76 kDa, with isoelectric points (*pI*) ranging from 4.64 to 7.78 (Table 1). Based on the *pI* value, the majority of proteins (28 members, 84.84%) were acidic.

**Table 1.** Features of Brassica LysM-RLK proteins.

Gene Name	Gene Stable ID	Chromosome / Scaffold Name	Gene Start (bp)	Gene End (bp)	Strand	Length	Weight (kDa)	pI	Localization
BnLYK1	GSBRNA2T00068250001	A4	11744720	11746625	−1	602	65.77	5.32	PlasmaMembrane
BnLYK2	GSBRNA2T00021233001	A6	1183129	1186599	1	633	69.26	5.95	PlasmaMembrane
BnLYK3	GSBRNA2T00085753001	A8	1474443	1478188	1	665	72.76	5.74	PlasmaMembrane
BnLYK4	GSBRNA2T00010694001	C3	6266494	6270198	−1	664	72.61	5.84	PlasmaMembrane
BnLYK5	GSBRNA2T00149947001	C4	37709852	37711907	−1	598	65.24	5.41	PlasmaMembrane
BnLYP1	GSBRNA2T00065733001	A2	11064901	11067027	−1	422	43.95	5.12	Extracellular
BnLYP2	GSBRNA2T00125469001	A6	8524836	8526824	1	414	43.22	4.83	PlasmaMembrane
BnLYP3	GSBRNA2T00147404001	A6	17944793	17946593	−1	362	38.19	7.78	Extracellular
BnLYP4	GSBRNA2T00047588001	A7	2800269	2802134	−1	364	38.76	7.31	Extracellular
BnLYP5	GSBRNA2T00102765001	A8	15737129	15739025	−1	415	43.35	4.64	PlasmaMembrane
BnLYP6	GSBRNA2T00125046001	C3	32553799	32555600	1	362	38.17	6.12	Extracellular
BnLYP7	GSBRNA2T00119287001	C5	10798534	10800145	1	416	43.47	4.83	Extracellular
BnLYP8	GSBRNA2T00145018001	C7	9946891	9949154	1	364	38.83	7.77	Extracellular
BnLYP9	GSBRNA2T00066647001	C8	22779362	22781576	1	414	43.35	4.64	PlasmaMembrane
BnLYP10	GSBRNA2T00080311001	Cnn	10720216	10722392	−1	422	4.08	5.12	Extracellular
BnLysMn1	GSBRNA2T00081535001	Ann	7838524	7839838	−1	260	28.92	6.22	Extracellular
BnLysMn2	GSBRNA2T00148567001	C4	16939318	16941025	1	260	28.91	6.22	PlasmaMembrane
BoLYK1	Bo3g181300	C3	62973513	62976832	1	665	72.66	5.76	PlasmaMembrane
BoLYK2	Bo4g151880	C4	41862968	41864776	−1	602	65.6	5.32	PlasmaMembrane
BoLYP1	Bo2g092750	C2	25211400	25213688	1	475	5	6.7	Extracellular
BoLYP2	Bo3g092250	C3	33860217	33861898	1	362	38.27	6.13	Extracellular
BoLYP3	Bo5g034850	C5	11423329	11424940	1	416	43.45	4.83	Extracellular
BoLYP4	Bo7g014960	C7	5770672	5773498	−1	364	38.86	7.77	Extracellular
BoLYP5	Bo8g071130	C8	23528006	23530169	1	414	43.31	4.72	PlasmaMembrane
BoLysMn	Bo4g078290	C4	17241540	17243011	−1	260	28.92	6.22	PlasmaMembrane
BrLYK1	Bra032146	A4	10968263	10970071	−1	602	65.84	5.4	PlasmaMembrane
BrLYK2	Bra018937	A6	1151733	1154752	1	634	69.2	6.04	PlasmaMembrane
BrLYP1	Bra008320	A2	13690581	13692699	1	422	43.95	5.12	Extracellular
BrLYP2	Bra017956	A6	8777590	8779193	1	414	43.25	4.83	Extracellular
BrLYP3	Bra009660	A6	17045973	17047648	−1	362	38.1	7.78	Extracellular
BrLYP4	Bra002021	A7	2492928	2494398	−1	364	38.84	6.68	Extracellular
BrLYP5	Bra016402	A8	17512432	17514340	−1	415	43.33	4.64	PlasmaMembrane
BrLysMn1	Bra038977	Scaffold000157	101880	103520	−1	260	28.91	6.52	Extracellular

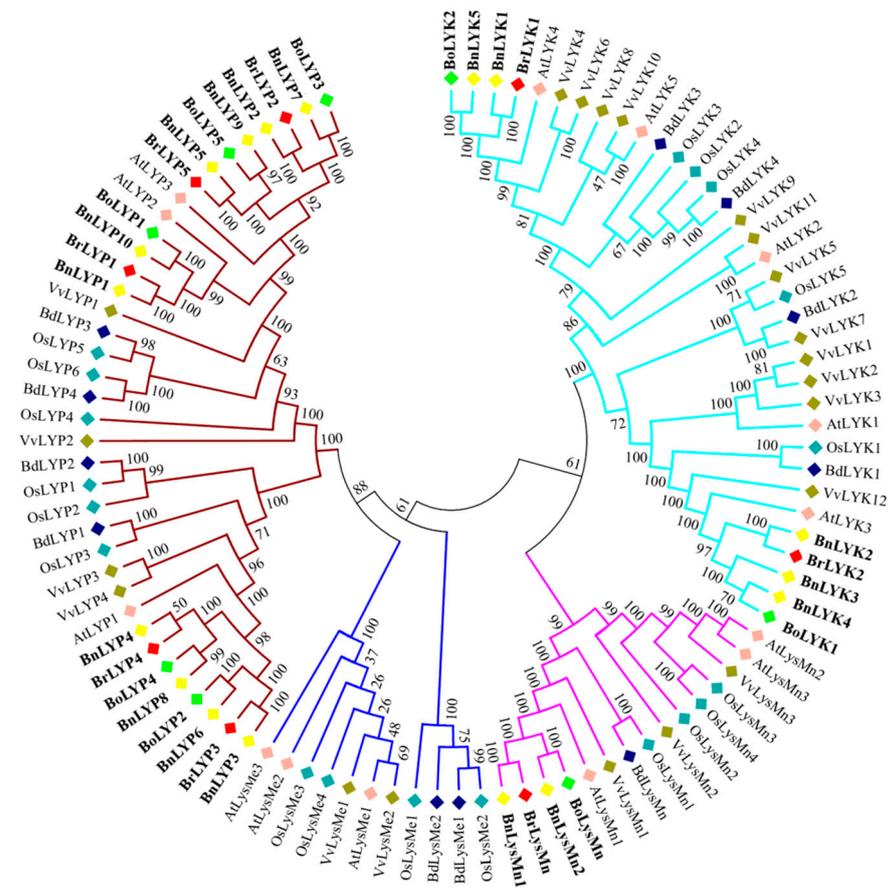
### 3.2. Phylogenetic Analysis of LysM-RLK Proteins

A neighbor-joining phylogenetic dendrogram was constructed to establish the link between *Brassica* LysM-RLK proteins and their homologous in other plants. According to Figure 1, the LysM-RLKs in *Brassica* were highly similar to their counterparts in Arabidopsis (*At*), rice (*Os*), and grapevine (*Vv*). LysM-RLK proteins were divided into four subfamilies: LYK, LYP, LysMe, and LysMn. Except for LysMn, all subfamilies have been identified in the three *Brassica* species investigated. Based on previous studies in Arabidopsis, AtLYP1 and AtLYP3 recognize peptidoglycan while AtCERK1, AtLYK4, and AtLYK5 recognize chitin. Therefore, due to the existence of the BnLYP2, BrLYP2, BoLYP2, BnLYP3, BrLYP3, BoLYP3, BnLYP4, BrLYP4, BoLYP4, BnLYP5, BrLYP5, BoLYP5, BnLYP6, BoLYP7, BnLYP8, BnLYP9 in the clade of AtLYP1 and AtLYP3, they can recognize peptidoglycan while BnLYK1, BrLYK1, BnLYK5, and BoLYK2 formed a monophyletic cluster with AtLYK4 confirming their ability to recognize chitin. It seems that LYP and LYK subgroups can specifically identify peptidoglycan and chitin, respectively. However, some studies reported that some members of the LYP subfamily can bind to both chitin and peptidoglycan ligands such as LYP1, LYP4, LYP5, and LYP6 [25,63,75,76].

### 3.3. Gene Duplication, Gene Location on the Chromosomes, and Selection Pressure of LysM-RLK Genes

The chromosomal distribution of 33 *Brassica* LysM-RLK was unequal on chromosomes (Figure 2). Chromosome (Chr) A6 in *B. rapa* and *B. napus* had the most genes, while ChrC3 and ChrC4 revealed the highest number of genes in *B. oleracea*. The genes of each subfamily were located on different chromosomes. In LYP subfamily of *B. napus*, LYP1, LYP2-3, LYP4, LYP5, LYP6, LYP7, LYP8, LYP9, and LYP10 were found on ChrA2, ChrA6, ChrA7, ChrA8, ChrC3, ChrC5, ChrC7, and ChrC8, respectively. BnLysMn1 and BnLysMn2 were identified on ChrAnn and ChrC4, respectively, in the LysMn subfamily of *B. napus*. Only segmental duplication was found in the *Brassica* LysM-RLK gene family, according to duplication analyses (Supplementary Materials: Table S2). Gene duplication is an effective phenomenon contributing to the abundance of duplicate genes in plant genomes which have contributed to the evolution of novel functions. To indicate selection pressure between

duplicated genes, the Ks, Ka, and Ka/Ks parameters were investigated for 43 paired genes (Supplementary Materials: Table S3).



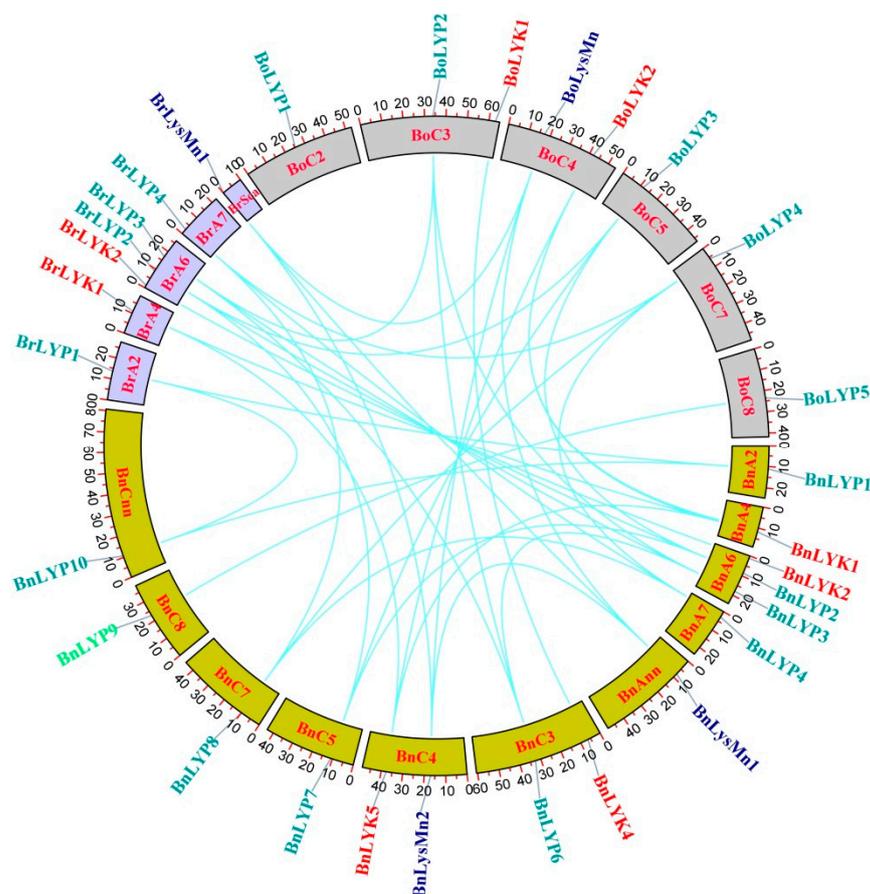
**Figure 1.** Phylogenetic relationships of *LysM*-RLK genes from *Brassica napus* (Bn), *Brassica rapa* (Br), *Brassica oleracea* (Bo), *Oryza sativa* (Os), *Arabidopsis thaliana* (At), *Vitis vinifera* (Vv), and *Brachypodium distachyon* (Bd). Colored branches have been used to depict various subfamilies. The phylogenetic dendrogram was constructed using MEGA 7 software and the neighbor-joining (NJ) method with 1000 bootstraps.

Except for *BnLysMn2/BnLysMn1*, *BnLysMn2/BrLysMn*, and *BnLYP6/BrLYP3*, the Ka/Ks ratio of 43 paired genes was less than 1, showing negative selection to maintain their function during *Brassica* evolution. The Ka/Ks ratio for three paired genes (*BnLysMn2/BnLysMn1*, *BnLysMn2/BrLysMn*, and *BnLYP6/BrLYP3*) was more than one, indicating positive selection, which resulted in their various functions as a result of mutations during their evolution.

### 3.4. Exon-Intron Structures and Conserved Motifs of *Brassica* *LysM*-RLKs

The MEME tool was used to find conserved motifs in *Brassica* *LysM*-RLK protein sequences (Supplementary Materials: Table S3). According to the data, 15 conserved motifs have been discovered, although the lowest number of motifs was detected in *LysMn* with 6 motifs (Figure 3A). The highest number of motifs was related to the *BoLYP1* with 13 motifs, followed by *BnLYP1*, *BrLYP1*, and *BnLYP10* with 12 motifs. As expected, each subgroup showed approximately similar motif compositions. *Brassica* *LysM*-RLK contains 0 to 10 introns, with *BoLYP4* being the longest intron, according to the exon-intron structural study (Figure 3B). Intron-free *Brassica* *LysM*-RLK genes account for 9.09% of the genome. The majority of *Brassica* *LysM*-RLK genes exhibited zero, one, or two forms of intron splicing, but *BnLYK5*, *BoLYK2*, *BoLYK1*, *BrLYK1*, *BnLysMn1*, *BnLysMn2*, *BoLysMn*, and *BrLysMn* had intron phase splicing zero. Exons ranged from one to five in *Brassica* *LysM*-RLKs, whereas *BnLYK2*, *BnLYK4*, *BnLYK3*, *BrLYK2*, and *BoLYK1* contained nine and eleven exons,

respectively. The highest amount of diversity in the number of exons was observed in the LYK subfamily, which indicates a selective pressure to obtain different functions during the evolution of *Brassica* [77]. Each subfamily showed similar intron splicing phases. The LysMn subfamily only displayed splicing phase zero, whereas the LYP subfamily showed all three splicing phases. Based on the splicing phase, the LYK subfamily was separated into two groups: (1) *BnLYK5*, *BoLYK2*, *BoLYK1*, and *BrLYK1* with splicing phase zero, and (2) *BoLYK1*, *BnLYK2*, *BrLYK2*, *BnLYK3*, and *BnLYK4* with all three splicing phases. The untranslated region was only found in 10 of the *Brassica* LysM-RLKs including *BnLYP2-3*, *BnLYP4*, *BnLYP6*, *BnLYP8-10*, *BnLYK1-5*, and *BnLysMn2*.

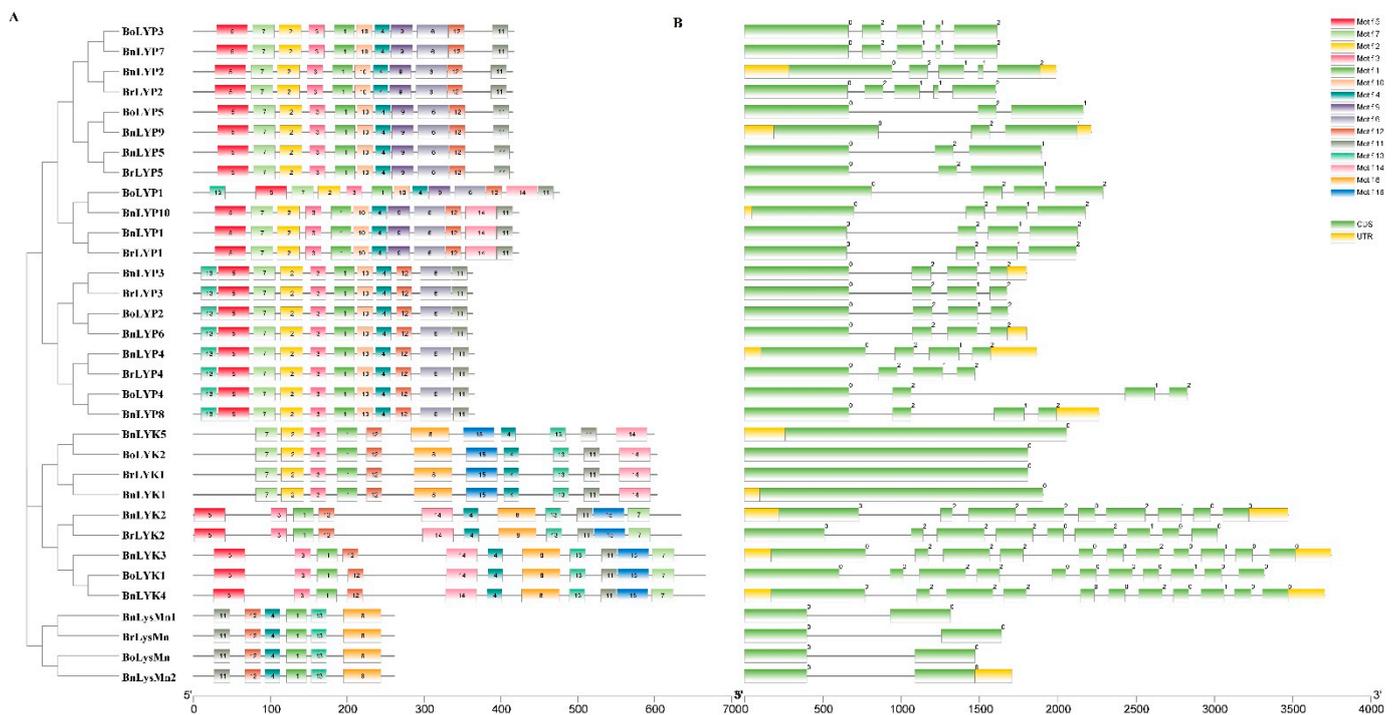


**Figure 2.** The chromosomal location of *LysM*-RLK genes and the duplication relationship between them. Colored boxes represent chromosomes. Curves are used to show gene duplications.

### 3.5. The Prediction of Cis-Regulatory Elements, Simple Sequence Repeats (SSR) Markers, and *Brassica* LysM-RLK-Targeted miRNAs

PlantCare was used to detect cis-regulatory elements in 1500 bp upstream of the *Brassica* LysM-RLK start codon (Supplementary Materials: Table S4). The *Brassica* LysM-RLK gene family has been discovered to have 70 cis-elements that can control gene expression in response to five different factors: environmental stresses, light, circadian, phytohormones, and developmental stages. The highest frequency of cis-acting elements in *B. napus*, *B. oleracea*, and *B. rapa* was related to ARE (94.11%), MYC (100%), and ARE (100%), respectively. The lowest frequency of cis-regulatory elements was related to GC-motif (only in *BoLYP2*) AT-rich sequence (only in *BnLYK4*), CARE (only in *BnLYK2*), GTGGC-motif (only in *BnLYR1*), MSA-like (only in *BoLYR2*), and F-box (only in *BnLYP5*). *Brassica* LysM-RLK contained 218 stress-responsive elements, indicating that they may have a role in regulating the *Brassica* response to different environmental challenges. 168, 161, and 75 cis-acting elements associated with phytohormones, light, and different tissues were also detected.

Therefore, *Brassica* LysM-RLKs have the potential to play a role in a variety of processes. 22 SSRs were identified in 16 out of 33 *Brassica* LysM-RLKs (13 SSRs in *B. napus*, 5 SSRs in *B. rapa*, and 4 SSRs in *B. oleracea*) (Table 2). Most genes had a single SSR except *BnLYP5* (2 SSRs), *BnLYP2*, and *BnLYP9* (4 SSRs each). The highest frequency was related to tetra-nucleotide repeats (9 SSRs) followed by di-nucleotide repeats (6 SSRs), tri-nucleotide repeats (4 SSRs), and pentanucleotide repeats (3 SSRs). 39 miRNAs for 12 *Brassica* LysM-RLKs targets have been detected (Supplementary Materials: Table S5). miRNAs and their targets did not have a one-to-one relationship, and many miRNAs shared a common target. For instance, 10 miRNAs named bra-miR156a-5p, bra-miR156b-5p, bra-miR156c-5p, bra-miR156d-5p, bra-miR156e-5p, bra-miR156f-5p, bra-miR156g-5p, bra-miR5725, bra-miR5721, and bra-miR9565-3p co-targeted *BrLYP2* transcript. One miRNA such as bna-miR390a can suppress the expression of multiple targets including *BnLYK1*, *BnLYK3*, *BnLYK4*, and *BnLYK5* as well.



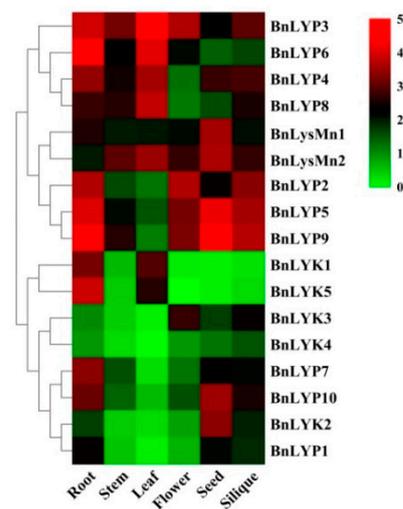
**Figure 3.** The conserved motifs (A) and exon-intron structure (B) of *LysM*-RLK genes in *Brassica* species. Exons and introns were represented by green boxes and black lines, respectively. Different motifs are shown by different colors. Exon-intron structure and Motifs were determined using gene structure display server (GSDS) and MEME online tool, respectively.

### 3.6. Expression Analysis of *BnLysM*-RLK Genes at Various Tissues under Biotic and Abiotic Stresses

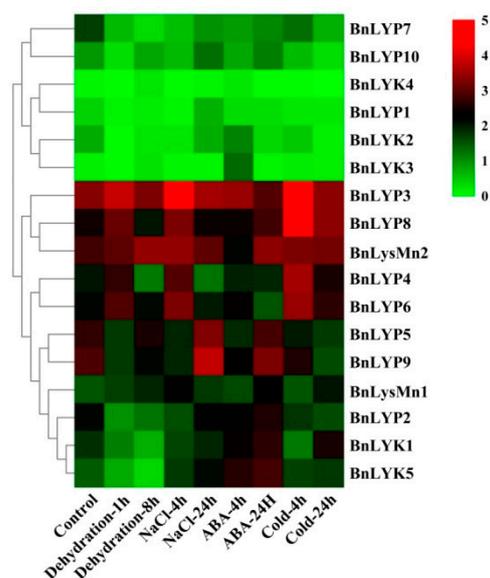
Because of its high content of unsaturated fatty acids and proteins, *B. napus* is considered one of the plants that produce the healthiest oils. Due to its outstanding properties, such as rapid growth, this plant is also used as a useful species for genetic and molecular studies of development and adaptation to diverse conditions. Therefore, in the current study, the expression of *LysM*-RLK genes has been investigated in *B. napus*. RNA-seq data sets for *B. napus* at different developmental stages tissues have been studied in leaf, flower, root, seed, stem, and silique to discover the related *LysM*-RLKs (Figure 4, Supplementary Materials: Table S6).

**Table 2.** Simple sequence repeats (SSR) were detected in Brassica LysM-RLK genes.

Seq ID	Count	Motif
BnLYP8	1	(CAAG)3
BnLYP4	1	(CAAG)3
BnLYK1	1	(CTC)4
BnLYP5	2	(CCTT)4, (TGTGG)3
BnLYP9	4	(CT)7, (CT)9, (AAG)4, (CCTT)4
BnLYP2	4	(TC)7, (TC)6, (GA)7, (AGTC)3
BrLYP4	1	(CAAG)3
BrLYP2	1	(AGTC)3
BrLYK1	1	(CTC)4
BrLYP5	1	(TGTGG)3
BrLysMn	1	(TATAT)3
BoLYP4	1	(CAAG)3
BoLYP1	1	(CT)9
BoLYK2	1	(CTC)4
BoLYP5	1	(CCTT)4

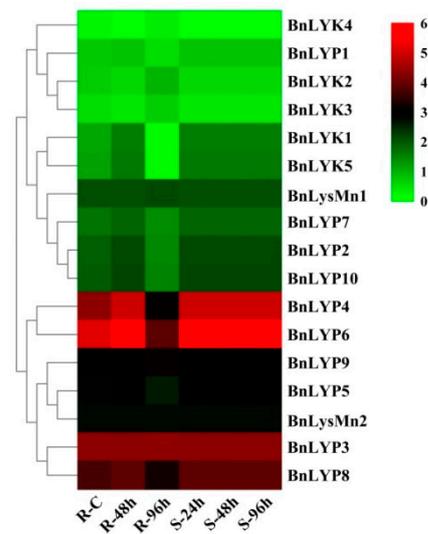
**Figure 4.** The expression pattern of *LysM-RLK* genes in different tissues. The color boxes indicate expression values, the lowest (green), medium (black), and the highest (red).

Different expression patterns have been observed in members of the *LysM-RLK* family. All members of the LYP subfamily revealed moderate to high transcript levels at all developmental stages and tissues except *BnLYP1* (low expression in all tissues except seed), *BnLYP2* (low expression in stem and leaf), *BnLYP6* (low expression in seed and silique), *BnLYP7* and *BnLYP10* (low expression in the stem, leaf, and flower), *BnLYP8* (low expression in flower and seed), and *BnLYP9* (low expression in leaf) and LYP4 low expression in flower. The highest expression in this subfamily was related to seed (*BnLYP8*, followed by *BnLYP5*), leaf (*BnLYP3*, followed by *BnLYP6*), silique (*BnLYP9*, followed by *BnLYP5*), flower (*BnLYP2*, followed by *BnLYP3*), and stem (*BnLYP3*). In the LYK subfamily, all *BnLYKs* demonstrated low expression except *BnLYK1* and *BnLYK5* (high level of transcripts in root and moderated expression in leaf), *BnLYK2* (high level of transcripts in seed), and *BnLYK3* (moderated expression in flower and silique). However, *BnLYK5* showed no obvious expression in the flower. Based on RNA-seq data analysis of the *BnLysMn* subfamily, all members demonstrated moderate to high levels of transcripts in tissues. The expression patterns of *BnLysM-RLK* genes have been examined to predict their role in responding to abiotic stresses as well (Figure 5, Supplementary Materials: Table S7).

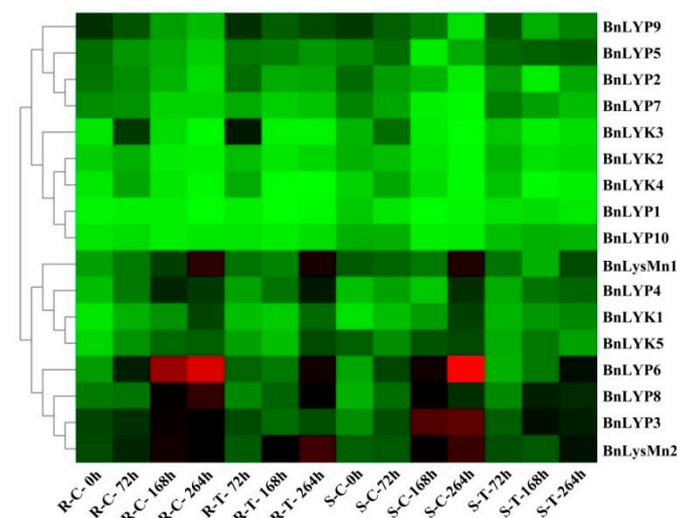


**Figure 5.** The expression pattern of *LysM-RLK* genes under abiotic stresses. The color boxes indicate expression values, the lowest (green), medium (black), and the highest (red).

In response to dehydration after one hour, the down-regulated expression has been observed in all *BnLysM-RLKs* while the expression of *BnLYP3-4*, *BnLYP6*, *BnLYP8*, and *BnLysMn1-2* was up-regulated. After 8 h of dehydration, the expression of all *BnLysM-RLKs* has been down-regulated obviously except *BnLYK3-4*, and *BnLysMn1-2* which showed up-regulation. The expression of *BnLysMn1-2* could be up-regulated by all the studied stresses except *BnLysMn1* and *BnLysMn2* with no obvious and down-regulated expression in response to cold and ABA after four hours, respectively. Under NaCl treatment, the expression of *BnLYK2-3*, *BnLYP4*, and *BnLYP6-8* has been decreased whereas the expression of other *BnLysM-RLKs* has been induced more significantly at 24 h. The expression of *BnLYP1* and *BnLYP7* has been suppressed by all the studied stresses except *BnLYP1* showed up-regulation in response to NaCl after 24 h. The expression of 9 *BnLysM-RLKs* genes has been up-regulated under ABA stress after four hours including *BnLYK1-5*, *BnLYP2-3*, *BnLYP6*, and *BnLysMn1* while the transcript level of the *BnLYK2-4*, *BnLYP1*, *BnLYP3-4*, and *BnLYP6-7* was down-regulated after 24 h of ABA treatment. After 24 h of cold stress, the expression of *BnLYK2*, *BnLYK4*, *BnLYP1-2*, *BnLYP5*, *BnLYP7*, and *BnLYP9-10* genes has been down-regulated. The RNA-seq data sets were applied for analyzing the expression of *BnLysM-RLKs* in response to fungal pathogens including *Leptosphaeria maculans* and *Sclerotinia sclerotiorum*. In response to *S. sclerotiorum*, *BnLYP3-6*, *BnLYP8-9*, and *BnLysMn2* revealed moderate to high expression in resistance, sensitivity, and control *B. napus* whereas the lowest expression was related to *BnLYK4*. As illustrated in Figure 6, *BnLYP3-4* and *BnLYP6* are consistently highly expressed in response to *S. sclerotiorum*, showing that these genes are likely involved in response to a fungal pathogen. The expression of all *BnLysM-RLKs* revealed down-regulated expression after *L. maculans* infection while *BnLY5* and *BnLY9* showed up-regulation (Figure 7). Similarly, *BnLYP7* and *BnLYK2* have been up-regulated in both susceptible and resistant cultivars except in resistant *B. napus* after 72 h of infection. *BnLYP6* has been suppressed by *L. maculans* infection in both susceptible and resistant cultivars. In general, the expression of *BnLysM-RLKs* in response to *S. sclerotiorum* infection was much higher than the response to *L. maculans* infection (Supplementary Materials: Tables S8 and S9).



**Figure 6.** The expression pattern of *LysM-RLK* genes in response to *Sclerotinia clerotiorum* infection. The color boxes indicate expression values, the lowest (green), medium (black), and the highest (red). R, S, and C indicate resistant (J964), susceptible (J902), and control plants, respectively.



**Figure 7.** The expression pattern of *LysM-RLK* genes in response to *Leptosphaeria maculans* infection. The color boxes indicate expression values, the lowest (green), medium (black), and the highest (red). R, S, C, and T indicate resistant (DF78), susceptible (Westar), control, and treatment plants, respectively.

### 3.7. BnLYP6 Structural Modeling and Docking Studies

In the current investigation, the highest expression in response to biotic stress was related to *BnLYP6*, thus, its molecular structure and ligand-enzyme interaction were investigated. Because PGN and chitin are structurally similar, LYP4 and LYP6 may also physically bind to chitin [75]. I-TASSER and ModRefinder servers have been used to predict and refine three-dimensional structures of *BnLYP6* protein. Based on the results of the Ramachandran analysis of non-refined and refined models, the residue count increased in favored regions from 63.7% to 78.1%, which indicates the efficiency of the refinement stage and increase the quality of the modeled structure (Supplementary Materials, Table S10 and Figure S1). The modeled structure for *BnLYP6* revealed 8 helices, 14 strands, 6 beta hairpins, 71 beta turns, and 4 gamma turns (Figure 8A). The *BnLYP6* structure contains two domains, including the LysM domain I (residues 113–159) and the LysM domain II (residues 177–220) (Figure 8B). The LysM domain is varied in size, ranging from 35 to 50 amino acids. LysM domain I and



### 3.8. The Codon Usage Bias Analysis of *Brassica LysM-RLK*

The results of the codon usage bias analysis have been shown in Supplementary Materials (Supplementary Materials, Table S11). The GC value for *Brassica LysM-RLK* genes was between 0.437 and 0.548, while the GC3s value was between 0.383 and 0.617. Because of the strong correlation between GC and GC3s, the mutation is the most important factor in codon creation (Table 3).

**Table 3.** The correlation coefficient between the parameters of codon usage of *Brassica LysM-RLK* gene family.

	CAI	CBI	Fop	ENC	GC3s
CBI	0.7 **				
Fop	0.73 **	0.99 **			
ENC	−0.12 <sup>ns</sup>	−0.03 <sup>ns</sup>	−0.1 <sup>ns</sup>		
GC3s	0.59 **	0.81 **	0.78 **	0.18 <sup>ns</sup>	
GC	0.69 **	0.88 **	0.85 **	0.25 <sup>ns</sup>	0.90 **

<sup>ns</sup> and \*\* are not-significant and significant at 1% probability level respectively.

The CAI (codon adaptation index), which was in the range of 0.221–0.262 in *Brassica LysM-RLKs*, is typically used to predict gene expression levels. The closer CAI is to 1, the stronger the codon preference and the higher the gene expression. A relative synonymous codon usage (RSCU) > 1 implies that codons are used more frequently than other synonymous, an RSCU = 1 indicates that codons are not preferred, and an RSCU of 1 indicates that codons are rarely utilized by genes [80]. There are 21 codons in *BnLYP7*, 22 *BnLYP9*, 23 codons in *BoLYP3-5*, *BnLYP4-5*, and *BnLYK4*, 24 codons in *BrLYP4-5* and *BnLYP2*, 25 codons in *BoLYK1-2*, *BoLYP1*, *BrLYP2*, *BrLYK1*, *BnLYP5*, and *BnLYK5*, 26 codons in *BnLYK3*, 27 codons in *BrLYP3*, *BrLysMn*, and *BnLYK1*, 28 codons in *BoLYP2*, *BnLYK2*, *BnLYP1*, *BnLysMn1*, and *BnLYP3*, 29 codons in *BrLYP1* and *BnLYP6*, and 30 codons in *BoLysMn*, *BrLYK2*, *BnLysMn2*, and *BnLYP1* with RSCU > 1 indicating that these are the most desired codons for each gene. The higher RSCU value (the more frequent codons for each gene) is shown in red, while the lower RSCU value is shown in blue (Figure 9). According to the RSCU value, *Brassica LysM-RLKs* were divided into four clusters: cluster I (*BnLYK2-3*, *BoLYK1-2*, *BrLYK1*, and *BnLYK4-5*), cluster II (*BnLysMn1-2*, *BrLysMn*, and *BoLysMn*), cluster III (*BrLyP1-2*, *BnLYP1-2*, *BoLYP1*, *BoLYP3*, *BoLYP5*, *BrLYP5*, *BnLYP5*, and *BnLYP7*), and cluster IV (*BoLYP2*, *BnLYP3-4*, *BrLYP3-4*, *BoLYP4*, *BnLYP6*, and *BnLYP8*). Each cluster had a similar preference for codons.



into LYK, LYP, and LysMe groups. The distribution of *LysM-RLK* was uneven *Brassica* genome. Most, if not all, flowering plants had one or more genome duplication events in their evolution [84]. In the current study, only segmental duplication resulted in multiple copies of *LysM-RLK* genes in *Brassica*.

The Ka/Ks ratios of the most duplicated *LysM-RLKs* were less than 1 except for three duplicated gene pairs (*BnLysMn2/BnLysMn1*, *BnLysMn2/BrLysMn*, and *BnLYP6/BrLYP3*) with Ka/Ks more than 1 and two duplicated gene pairs (*BnLYP7/BoLYP3* and *BnLYP8/BoLYP4*) with no Ka/Ks value due to the same sequence. It should be noted that during evolution, changes in the coding region of duplicated genes resulted in various functions due to amino acid substitution or exon-intron structural divergence [85]. Because of the high purifying selection in the *LysM-RLK* gene family, the importance of the functional role of *Brassica LysM-RLK* genes has been determined. According to the phylogenetic tree, it was shown that *Brassica LysM-RLKs* have a close relationship with their counterparts due to their sequence conservation and similar function. The amino acid compositions of each cluster were similar, implying that the phylogenetic distribution of *Brassica LysM-RLK* proteins is associated with their motif contents. All members of the LYK cluster contained 10 common motifs 1, 3-4, 7-8, and 11-15. The difference between the members of this subfamily was related to motif 5 in the clade of BoLYK1, BnLYK2, BrLYK2, BnLYK3, and BnLYK4, while the clade of BnLYK1, BrLYK1, BoLYK2, and BnLYK5 contained motif 2. These results are completely consistent with the results of the phylogenetic tree. The LysMn subfamily had common motifs 1, 4, 8, 12, and 13. In the LYP subfamily, three groups were observed. The first group (Cluster I) consisted of BrLYP2, BnLYP2, BoLYP3, BoLYP5, BrLYP, BnLYP5, BnLYP7, and BnLYP9 proteins with common motifs 1-7 and 9-12. The second group (Cluster II), including BnLYP1, BoLYP1, BrLYP1, BnLYP10 with common motifs 1-7, 9-12, and 14 except for BoLYP1 with extra specific motif 13. The third group (Cluster III) consisted of BoLYP2, BrLYP3, BnLYP3, BnLYP4, BrLYP4, BoLYP4, BnLYP6, and BnLYP8 demonstrated 11 same motifs 1-7 and 10-13. The difference between clusters I and II was related to the existence of motifs 14 in the second cluster while cluster III was separated from the above two clusters due to the lack of motif 9. The structure of exons and introns, as well as the splicing phase, play crucial roles in the evolution of gene families [86]. The high and highest conservations were found in intron phases 0 and 1, respectively, while the lowest conservation was found in intron phase 2 [87,88]. The frequency of phases 0 and 1 in all subfamilies was higher than in phase 2, including LYK (63.63%), LYP (55%), and LysMn (100%) indicating high conservation of protein function during *Brassica* evolution. 9.09% of *Brassica LysM-RLK* genes were intronless.

The study of promoter regions is necessary to understand the function of *Brassica LysM-RLK* genes. In response to environmental stresses, transcription factors play a significant role. They bind to the target genes' promoters, regulating their expression [89]. The presence of regulatory components related to stress, developmental stage, light, and phytohormones suggests that *LysM-RLKs* have a role in the plant's response to a variety of biological processes. Several cis-elements associated to plant resistance against biotic and abiotic stresses were identified based on the promoter analysis, including ARE, DRE, GC-motif, LTR, MBS, MYB, MYC, STRE, AP1, S-box, W-box, WUN-motif, and WRE3. The TGACG and CGTCA motifs are found on methyl jasmonate-responsive genes [90]. Senescence, seed germination, and response to biotic and abiotic stressors are all affected by jasmonate as well [91]. In response to ABA, the ABRE, ABRE3a, and ABRE4 motifs are activated, resulting in drought and salinity tolerance in plants. The high frequency of cis-acting elements associated with response to drought, pathogen, cold, ABA, auxin, jasmonate, gibberellin, and ethylene suggests that *LysM-RLK* genes are active in a variety of stress responses in *Brassica* species. However, the existence of specific regulatory elements is not sufficient evidence for these genes' responses to specific hormones or stresses, requiring the use of laboratory procedures to precisely determine their function. SSRs are 1-6 nucleotide tandem repeats that have been shown to play a crucial function in gene regulation [92]. In the current study, tetra-nucleotide repeats (40.91%) were found to be more common than

other SSRs. The type of dominant SSRs varies in various plant species, and the abundance of AT repeats is higher in the dicots genome than monocots [93]. SSR polymorphisms in *LysM-RLK* may be examined in different cultivars in the future, and they may be useful for marker-assisted selection (MAS) development in *Brassica* genetic improvement to choose genotypes with higher resistance to various stresses. MicroRNAs (miRNA) are non-coding small RNAs with a length of 19–24 bp. They are crucial in the regulation of post-transcriptional modifications. Plants, animals, and viruses all have miRNAs. Plant development and responses to environmental stressors are also influenced by them [94]. *Brassica* miRNAs targeted 6, 5, and 1 transcript in the LYK (*BnLYK1*, *BrLYK1*, *BoLYK1*, *BnLYK3-5*), LYP (*BnLYP2-3* and *BrLYP2-3*), and LysMn (*BnLysMn2*) subfamilies, respectively. No *LysM-RLK*-targeted miRNA was found in LysMn and LYP subfamilies. miR156 is required for the vegetative phase transition of a plant from a juvenile to an adult [95]. Under normal growth conditions, auxin-induced miR390 stimulates lateral root development [96]. Therefore, *BnLYK1* and *BnLYK3-5* are likely to play a role in root development. miR396 with reduced activity has been demonstrated to give widespread resistance to necrotrophic and hemibiotrophic fungal infections in *Arabidopsis* [97], thus, *BnLysMn2* may be involved in the *B.napus* defense against fungal infections. miR397 has been reported that target laccase family genes through transcript cleavage in *Arabidopsis* and rice [98]. As a result, they are required for the maintenance of cell walls and vascular integrity, implying that they play a role in plant defense against various stresses [99]. In *Arabidopsis*, banana, and rice, miR397 has been shown to have a major impact on plant biomass and yield [99–101] that targets *BnLYK4* in this study. miR5717 regulates genes involved in lipid metabolism and pollen tube growth [102]. Therefore, *BrLYP3* is likely to have a role in reproductive development. It was hypothesized that miR5721 may target genes that encode biotinyl-lipoyl-containing proteins [103]. In *B. napus*, miR2111 plays a significant role in the response to phosphorus deficiency [104]. Finally, miR6029 has been reported to regulate fatty acid production during the development of *B. napus* seeds [105].

The expression profile of genes provides important information about the function of the genes that have been found. According to recent studies, *RLKs* are thought to play a crucial role in stress responses [106,107]. The highest number of *BnLysM-RLK* genes with moderate to high expression was observed in seeds (76.92%) followed by roots (76.47%), and silique (52.94%) while the lowest number of moderate to highly expressed genes was related to stem (35.39%) preceded by leaf and flower (41.17% each). *BnLYK5* was considered not expressed in flower tissue. The highest expression in root and flower tissues was related to *BnLYP6* and *BnLYP2*, respectively, while in stem and leaves the highest expression was related to *BnLYP3* and in seed and silique was related to *BnLYP9*. Most of the low-expression and high-expression genes were related to LYK and LYP subfamilies, respectively. The results demonstrated that the expression patterns of genes belonging to the same subfamily can differ significantly. For instance, the *BnLYP3* and *BnLYP8* of the LYP subfamily are consistently expressed at high levels while other LYP genes demonstrated a minimum expression except for *BnLYP4-6* and *BnLYP9* with moderate to high expressions. The results reinforced the hypothesis of divergence that the duplicated genes may be the result of one of two processes: 1) subfunctionalization, and 2) neofunctionalization. In the subfunctionalization process, some of the characteristics of new genes vary from the parental genes [108], whereas the new gene plays a different role in the neofunctionalization process due to differences in amino acid content [109]. Drought is one of the important environmental stresses that have negative effects on plant growth. *RLKs'* response to drought stress is influenced by ABA [107]. ABA is a key plant hormone, regulating the expression of genes involved in drought, salt, and osmotic stress responses [110]. As an ABA-dependent pathway, *Arabidopsis* receptor dead kinase 1 (RDK1) plays an important role in drought stress response. The *Arabidopsis* *rdk1* mutants were hypersensitive to drought stress due to the down-regulation of ABA-responsive genes [111]. Considering the present study, the expression of *LysM-RLK* genes in response to abiotic stresses varies depending on the stress type and duration. Thus, *BnLYP3* and *BnLYP8* genes were up-

regulated by salt after 4 h of treatment, while they were down-regulated after 24 h under salinity condition. The highest transcript level under dehydration conditions after 1 and 8 h was related to *BnLYP3* and *BnLysMN2*, respectively. Interestingly, in all treatments including salt (after 4 h), ABA (after 4 h), and cold (after 4 and 24 h) *BnLYP3* showed the highest expression while the transcript levels of *BnLYP9* and *BnLysMn2* was higher than other *LysM-RLKs* in response to salinity and ABA treatments after 24 h, respectively. These findings suggested that the *BnLYP3* gene may play a critical role in *B. napus* response to abiotic stresses, which can be utilized to improve the resistance of *B. napus* cultivars in future researches. We can also suggest this gene as a marker of abiotic stresses in *B. napus*. Pathogens and pests are believed to be capable of causing 50–60% losses in Brassica crop yield and quality, resulting in significant economic losses [112]. Sclerotinia stem rot is one of the most destructive diseases for *B. napus*, caused by *S. sclerotiorum*. The highest expression in response to *S. sclerotiorum* was related to *BnLYP6*, followed by *BnLYP4*, *BnLYP3*, *BnLYP8*, and *BnLYP5* in both susceptible and resistant cultivars except in resistant cultivar after 96 h that the highest expression was related to *BnLYP3*, followed by *BnLYP6*, *BnLYP8*, *BnLYP9*, and *BnLYP4*. Based on the results of Brotman et al. (2012), CERK1 (*LysM-RLK1*) receptor is required for chitinase-induced salt and heavy metal tolerance in plants. Likewise, they suggested that ectopic chitinases are largely involved in inducing plant immune response against pathogens mediated by the CERK1 receptor [26]. June et al. (2015) revealed that GbRLK plays an important role in modulating a variety of plant-pathogen interactions in *Gossypium barbadense*. According to their findings, the majority of the up-regulated genes associated with disease resistance were chitin responsive, implying that the transgenic Arabidopsis showed improved resistance against *Verticillium dahlia* by modulating the chitin response signaling pathway [113]. Blackleg disease, caused by *L. maculans*, is a serious production limitation in *B. napus*. It has been observed in all canola-growing regions except China and causes yearly yield losses of 10–20% [114]. Expression in *BnLysM-RLKs* is suppressed after *L. maculans* infection, except in *BnLYP5* and *BnLYP9* that their expression was slightly increased after pathogen infection. Taken collectively, all members of the gene family are expressed in *B. napus*.

The study of the *LysM-RLK* gene family in other plants also shows the response of these genes to fungal and bacterial pathogens. A study of transcriptome data has shown that the expression of wheat *LysM-RLK* genes is induced in response to Flg22 and chitin. Therefore, these genes are involved in wheat resistance to fungal and bacterial pathogens [40]. In *Citrus sinensis*, the expression of *LYK* genes have increased in response to *Xanthomonas citri*, the Citrus bacterial canker (CBC) causing plant bacterial pathogen, and the salicylic acid (SA), methyl jasmonate (MeJA), and abscisic acid (ABA) hormones. Accordingly, there is a link between the *LYK* genes, the ABA, SA, and MeJA signaling pathways, and CBC resistance [17]. *Fusarium graminearum* (Fg), the causative agent of *Fusarium* head blight (FHB), induces the expression of *BdLYK2*, *BdLYK3*, and *BdLYK4* genes in *Brachypodium distachyon*. On the other hand, the expression of *BdLYP1* and *BdLYP4* genes has decreased in response to this pathogen. The function of these genes seems to be similar to the Arabidopsis *AtLYP2* and *AtLYP3* genes, which are involved in responding to bacterial pathogens [78]. Although these results can confirm the specificity of *LYP* genes to bacterial PGN, in rice, *LYP4* and *LYP6* genes are dual-functional and can respond simultaneously to fungal chitin and PGN [115]. The present study also showed that *BnLYP6* gene expression is induced in response to fungal pathogens. On the other hand, molecular docking analysis showed that *BnLYP6* has a high affinity for chitin, which indicates the role of this gene in responding to fungal pathogens in *B. napus*. These results together indicate the different functions of *LysM-RLK* genes in the response of plants to bacterial and fungal pathogens as well as abiotic stresses [116].

CUB can represent the origin of a gene and can be utilized as a theoretical model for analyzing gene evolution and function [117]. The amount of ENC varies between 20 to 61, and the higher the ENC value, the weaker the CUB. The ENC of the *Brassica* *LysM-RLKs* ranged from 48.97 to 59.64, indicating that the codons of this family are not affected by

strong codon bias and there are various synonymous codons [118]. The CAI index varies from 0 to 1, which is typically applied to measure expression levels [119]. According to the CAI index (0.221–0.2621), the expression efficiency of the *BnLysM-RLKs* is almost low. Although the codon preference of highly expressed genes is stronger with a higher CAI and lower NC values, low-expression genes have more rare codons, resulting in a lower CAI and a higher NC. For instance, *BnLYP6* showed increased expression in response to biotic stresses with almost larger CAI and relatively lower NC. The optimal codon frequency is represented by the FOP and CBI indices, which range from 0 to 1 and -1 to 1, respectively. Based on the results of the FOP and CBI, the frequency of optimum codons in this gene family was low. The majority of *Brassica* LysM-RLKs showed a GC content of more than 0.5, implying that *Brassica* LysM-RLKs have obvious preference for GC. 69.69% of *Brassica* LysM-RLKs demonstrated a GC3s value greater than 0.5, indicating that G/C end codons are preferred.

## 5. Conclusions

Bioinformatic analyses were performed in this work to discover 33 *LysM-RLK* genes with significant structural diversity in three *Brassica* species. Based on the phylogenetic analysis, *Brassica* LysM-RLK genes were divided into three groups including LYK, LYP, and LysMn. Only segmental duplication was found during the investigation of the mechanism of gene family expansion. The function of most duplicated *Brassica* LysM-RLK genes has been conserved over evolution due to negative selection. During promoter analysis, several elements in the *Brassica* LysM-RLK promoters were found, showing that they play a role in stress response and plant growth. 22 SSR and 39 miRNA were detected which can be employed in MAS and genetic transformation, respectively. The functional involvement of *LysM-RLK* genes in *Brassica* tissues in response to environmental stressors was revealed by their expression patterns in diverse tissues. Due to the high expression of *BnLYP3* genes in response to *Sclerotinia* stem rot infection and *BnLYP3* in response to abiotic stresses, these genes can be exploited in the production of *B. napus* plants resistant to biotic and abiotic stresses. The discovery of these residues might be important in future investigations to improve the efficiency of the LYP6 enzymes and generate pathogen-resistant *B. napus* by site-directed mutagenesis. This research has given fundamental information on the *LysM-RLK* genes, which will be useful in future investigations aimed at improving Brassica quality.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells11010037/s1>, Figure S1: Ramachandran plot for non-refined and refined modeled structure. Table S1: Features of *Brassica* LysM-RLK proteins. Table S2: *Brassica* LysM-RLK domains. Table S3: Ka/Ks analysis of the *Brassica* LysM-RLK duplicated paired genes. Table S4: Putative conserved motifs of *Brassica* LysM-RLK. Table S5: List of Cis-acting regulatory elements in the *Brassica* LysM-RLK promoter. Table S6: Putative *Brassica* LysM-RLK-targeted miRNA. Table S7: RNA-seq data analysis of *Brassica* LysM-RLK at various tissues. Table S8: RNA-seq data analysis of *Brassica* LysM-RLK under abiotic stress. Table S9: RNA-seq data analysis of *Brassica* LysM-RLK in response to *Sclerotinia sclerotiorum* infection. Table S10: RNA-seq data analysis of *Brassica* LysM-RLK in response to *Leptosphaeria maculans* infection. Table S11: Relative synonymous codon usage analysis (RSCU) of *Brassica* LysM-RLKs.

**Author Contributions:** Conceptualization, A.A. and Z.H.; methodology, A.A.; software, A.A. and Z.H.; validation, C.J., Q.E. and A.A.; formal analysis, M.K.; investigation, Z.H. and A.A.; resources, A.A.; data curation, Z.H. and A.A.; writing—original draft preparation, Z.H.; writing—review and editing, Q.E., C.J. and M.K.; visualization, A.A.; supervision, C.J. and M.K.; project administration, M.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available as Supplementary Files.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Chisholm, S.T.; Coaker, G.; Day, B.; Staskawicz, B.J. Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell* **2006**, *124*, 803–814. [[CrossRef](#)] [[PubMed](#)]
- Nürnberger, T.; Kemmerling, B. Receptor protein kinases—pattern recognition receptors in plant immunity. *Trends Plant Sci.* **2006**, *11*, 519–522. [[CrossRef](#)] [[PubMed](#)]
- Lu, Y.; Tsuda, K. Intimate association of prr-and nlr-mediated signaling in plant immunity. *Plant Microbe Interact.* **2021**, *34*, 3–14. [[CrossRef](#)] [[PubMed](#)]
- Tirnaz, S.; Bayer, P.E.; Inturrisi, F.; Zhang, F.; Yang, H.; Dolatabadian, A.; Neik, T.X.; Severn-Ellis, A.; Patel, D.A.; Ibrahim, M.I. Resistance gene analogs in the Brassicaceae: Identification, characterization, distribution, and evolution. *Plant Physiol.* **2020**, *184*, 909–922. [[CrossRef](#)] [[PubMed](#)]
- Li, B.; Lu, D.; Shan, L. Ubiquitination of pattern recognition receptors in plant innate immunity. *Mol. Plant Pathol.* **2014**, *15*, 737–746. [[CrossRef](#)]
- Zipfel, C. Plant pattern-recognition receptors. *Trends Immunol.* **2014**, *35*, 345–351. [[CrossRef](#)]
- Yu, X.; Feng, B.; He, P.; Shan, L. From chaos to harmony: Responses and signaling upon microbial pattern recognition. *Annu. Rev. Phytopathol.* **2017**, *55*, 109–137. [[CrossRef](#)]
- Antolín-Llovera, M.; Ried, M.K.; Binder, A.; Parniske, M. Receptor kinase signaling pathways in plant-microbe interactions. *Annu. Rev. Phytopathol.* **2012**, *50*, 451–473. [[CrossRef](#)]
- Jose, J.; Ghantasala, S.; Roy Choudhury, S. Arabidopsis transmembrane receptor-like kinases (rlks): A bridge between extracellular signal and intracellular regulatory machinery. *Int. J. Mol. Sci.* **2020**, *21*, 4000. [[CrossRef](#)]
- Dievart, A.; Gottin, C.; Périn, C.; Ranwez, V.; Chantret, N. Origin and diversity of plant receptor-like kinases. *Annu. Rev. Plant Biol.* **2020**, *71*, 131–156. [[CrossRef](#)]
- Dubey, M.; Véléz, H.; Broberg, M.; Jensen, D.F.; Karlsson, M. Lysm proteins regulate fungal development and contribute to hyphal protection and biocontrol traits in *Clonostachys rosea*. *Front. Microbiol.* **2020**, *11*, 679. [[CrossRef](#)] [[PubMed](#)]
- Cui, H.; Tsuda, K.; Parker, J.E. Effector-triggered immunity: From pathogen perception to robust defense. *Annu. Rev. Plant Biol.* **2015**, *66*, 487–511. [[CrossRef](#)] [[PubMed](#)]
- Buist, G.; Steen, A.; Kok, J.; Kuipers, O.P. Lysm, a widely distributed protein motif for binding to (peptido) glycans. *Mol. Microbiol.* **2008**, *68*, 838–847. [[CrossRef](#)]
- Dworkin, J. Detection of fungal and bacterial carbohydrates: Do the similar structures of chitin and peptidoglycan play a role in immune dysfunction? *PLoS Pathog.* **2018**, *14*, e1007271. [[CrossRef](#)]
- Kaku, H.; Nishizawa, Y.; Ishii-Minami, N.; Akimoto-Tomiyama, C.; Dohmae, N.; Takio, K.; Minami, E.; Shibuya, N. Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11086–11091. [[CrossRef](#)] [[PubMed](#)]
- Zhang, X.-C.; Cannon, S.B.; Stacey, G. Evolutionary genomics of *Lysm* genes in land plants. *BMC Evol. Biol.* **2009**, *9*, 1–13. [[CrossRef](#)]
- Li, Q.; Qi, J.; Qin, X.; Hu, A.; Fu, Y.; Chen, S.; He, Y. Systematic identification of lysin-motif receptor-like kinases (*Lyks*) in *Citrus sinensis*, and analysis of their inducible involvements in citrus bacterial canker and phytohormone signaling. *Sci. Hortic.* **2021**, *276*, 109755. [[CrossRef](#)]
- Lehti-Shiu, M.D.; Shiu, S.-H. Diversity, classification and function of the plant protein kinase superfamily. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 2619–2639. [[CrossRef](#)]
- Zhang, X.-C.; Wu, X.; Findley, S.; Wan, J.; Libault, M.; Nguyen, H.T.; Cannon, S.B.; Stacey, G. Molecular evolution of lysin motif-type receptor-like kinases in plants. *Plant Physiol.* **2007**, *144*, 623–636. [[CrossRef](#)]
- Lee, W.-S.; Rudd, J.J.; Hammond-Kosack, K.E.; Kanyuka, K. *Mycosphaerella graminicola* lysm effector-mediated stealth pathogenesis subverts recognition through both *cerk1* and *cebip* homologues in wheat. *Mol. Plant Microbe Interact.* **2014**, *27*, 236–243. [[CrossRef](#)]
- Tanaka, S.; Ichikawa, A.; Yamada, K.; Tsuji, G.; Nishiuchi, T.; Mori, M.; Koga, H.; Nishizawa, Y.; O’Connell, R.; Kubo, Y. *HvCEBiP*, a gene homologous to rice chitin receptor *CEBiP*, contributes to basal resistance of barley to *Magnaporthe oryzae*. *BMC Plant Biol.* **2010**, *10*, 1–11. [[CrossRef](#)] [[PubMed](#)]
- Hayafune, M.; Berisio, R.; Marchetti, R.; Silipo, A.; Kayama, M.; Desaki, Y.; Arima, S.; Squeglia, F.; Ruggiero, A.; Tokuyasu, K. Chitin-induced activation of immune signaling by the rice receptor *CEBiP* relies on a unique sandwich-type dimerization. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E404–E413. [[CrossRef](#)]
- Kouzai, Y.; Mochizuki, S.; Nakajima, K.; Desaki, Y.; Hayafune, M.; Miyazaki, H.; Yokotani, N.; Ozawa, K.; Minami, E.; Kaku, H. Targeted gene disruption of *OsCERK1* reveals its indispensable role in chitin perception and involvement in the peptidoglycan response and immunity in rice. *Mol. Plant Microbe Interact.* **2014**, *27*, 975–982. [[CrossRef](#)] [[PubMed](#)]
- Kouzai, Y.; Nakajima, K.; Hayafune, M.; Ozawa, K.; Kaku, H.; Shibuya, N.; Minami, E.; Nishizawa, Y. *CEBiP* is the major chitin oligomer-binding protein in rice and plays a main role in the perception of chitin oligomers. *Plant Mol. Biol.* **2014**, *84*, 519–528. [[CrossRef](#)]

25. Petutschnig, E.K.; Jones, A.M.; Serazetdinova, L.; Lipka, U.; Lipka, V. The lysin motif receptor-like kinase (LysM-RLK) CERK1 is a major chitin-binding protein in *Arabidopsis thaliana* and subject to chitin-induced phosphorylation. *J. Biol. Chem.* **2010**, *285*, 28902–28911. [[CrossRef](#)]
26. Brotman, Y.; Landau, U.; Pnini, S.; Lisec, J.; Balazadeh, S.; Mueller-Roeber, B.; Zilberstein, A.; Willmitzer, L.; Chet, I.; Viterbo, A. The LysM receptor-like kinase LysM RLK1 is required to activate defense and abiotic-stress responses induced by overexpression of fungal chitinases in *Arabidopsis* plants. *Mol. Plant.* **2012**, *5*, 1113–1124. [[CrossRef](#)] [[PubMed](#)]
27. Espinoza, C.; Liang, Y.; Stacey, G. Chitin receptor CERK1 links salt stress and chitin-triggered innate immunity in *Arabidopsis*. *Plant J.* **2017**, *89*, 984–995. [[CrossRef](#)]
28. Wan, J.; Tanaka, K.; Zhang, X.-C.; Son, G.H.; Brechenmacher, L.; Nguyen, T.H.N.; Stacey, G. LYK4, a lysin motif receptor-like kinase, is important for chitin signaling and plant innate immunity in *Arabidopsis*. *Plant Physiol.* **2012**, *160*, 396–406. [[CrossRef](#)]
29. Willmann, R.; Lajunen, H.M.; Erbs, G.; Newman, M.-A.; Kolb, D.; Tsuda, K.; Katagiri, F.; Fliegmann, J.; Bono, J.-J.; Cullimore, J.V. *Arabidopsis* lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19824–19829. [[CrossRef](#)]
30. Faulkner, C.; Petutschnig, E.; Benitez-Alfonso, Y.; Beck, M.; Robatzek, S.; Lipka, V.; Maule, A.J. LYM2-dependent chitin perception limits molecular flux via plasmodesmata. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9166–9170. [[CrossRef](#)]
31. Narusaka, Y.; Shinya, T.; Narusaka, M.; Motoyama, N.; Shimada, H.; Murakami, K.; Shibuya, N. Presence of LYM2 dependent but CERK1 independent disease resistance in *Arabidopsis*. *Plant Signal. Behav.* **2013**, *8*, e25345. [[CrossRef](#)] [[PubMed](#)]
32. Qiu, W.; Feechan, A.; Dry, I. Current understanding of grapevine defense mechanisms against the biotrophic fungus (*Erysiphe necator*), the causal agent of powdery mildew disease. *Hortic. Res.* **2015**, *2*, 1–9. [[CrossRef](#)] [[PubMed](#)]
33. Leppyanen, I.V.; Pavlova, O.A.; Vashurina, M.A.; Bovin, A.D.; Dolgikh, A.V.; Shtark, O.Y.; Sendersky, I.V.; Dolgikh, V.V.; Tikhonovich, I.A.; Dolgikh, E.A. LysM receptor-like kinase lyk9 of *Pisum Sativum* L. may regulate plant responses to chitoooligosaccharides differing in structure. *Int. J. Mol. Sci.* **2021**, *22*, 711. [[CrossRef](#)] [[PubMed](#)]
34. Buendia, L.; Wang, T.; Girardin, A.; Lefebvre, B. The LysM receptor-like kinase SILYK10 regulates the arbuscularmycorrhizal symbiosis in tomato. *New Phytol.* **2016**, *210*, 184–195. [[CrossRef](#)] [[PubMed](#)]
35. Liao, D.; Sun, X.; Wang, N.; Song, F.; Liang, Y. Tomato LysM receptor-like kinase SILYK12 is involved in arbuscular mycorrhizal symbiosis. *Front. Plant Sci.* **2018**, *9*, 1004. [[CrossRef](#)] [[PubMed](#)]
36. Neik, T.X.; Amas, J.; Barbetti, M.; Edwards, D.; Batley, J. Understanding host–pathogen interactions in *Brassica napus* in the omics era. *Plants* **2020**, *9*, 1336. [[CrossRef](#)]
37. Poveda, J.; Francisco, M.; Cartea, M.E.; Velasco, P. Development of transgenic *Brassica* crops against biotic stresses caused by pathogens and arthropod pests. *Plants* **2020**, *9*, 1664. [[CrossRef](#)]
38. Zhang, X.; Lu, G.; Long, W.; Zou, X.; Li, F.; Nishio, T. Recent progress in drought and salt tolerance studies in *Brassica* crops. *Breed. Sci.* **2014**, *64*, 60–73. [[CrossRef](#)]
39. Chen, Q.; Li, Q.; Qiao, X.; Yin, H.; Zhang, S. Genome-wide identification of lysin motif containing protein family genes in eight rosaceae species, and expression analysis in response to pathogenic fungus *Botryosphaeria dothidea* in Chinese white pear. *BMC Genom.* **2020**, *21*, 1–20. [[CrossRef](#)]
40. Chen, Z.; Shen, Z.; Zhao, D.; Xu, L.; Zhang, L.; Zou, Q. Genome-wide analysis of LysM-containing gene family in wheat: Structural and phylogenetic analysis during development and defense. *Genes* **2021**, *12*, 31. [[CrossRef](#)]
41. Lehti-Shiu, M.D.; Zou, C.; Hanada, K.; Shiu, S.-H. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **2009**, *150*, 12–26. [[CrossRef](#)]
42. Lohmann, G.V.; Shimoda, Y.; Nielsen, M.W.; Jørgensen, F.G.; Grossmann, C.; Sandal, N.; Sørensen, K.; Thirup, S.; Madsen, L.H.; Tabata, S. Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. *Mol. Plant Microbe Interact.* **2010**, *23*, 510–521. [[CrossRef](#)] [[PubMed](#)]
43. Nazarian-Firouzabadi, F.; Joshi, S.; Xue, H.; Kushalappa, A.C. Genome-wide in silico identification of LysM-RLK genes in potato (*Solanum tuberosum* L.). *Mol. Biol. Rep.* **2019**, *46*, 5005–5017. [[CrossRef](#)] [[PubMed](#)]
44. Shiu, S.-H.; Karlowski, W.M.; Pan, R.; Tzeng, Y.-H.; Mayer, K.F.; Li, W.-H. Comparative analysis of the Receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* **2004**, *16*, 1220–1234. [[CrossRef](#)] [[PubMed](#)]
45. Bateman, A.; Coin, L.; Durbin, R.; Finn, R.D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E.L. Pfam: The protein families database. *Nucleic Acids Res.* **2004**, *32*, D138–D141. [[CrossRef](#)] [[PubMed](#)]
46. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* **2011**, *39*, W29–W37. [[CrossRef](#)] [[PubMed](#)]
47. Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; De Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **2012**, *40*, W597–W603. [[CrossRef](#)]
48. Almagro Armenteros, J.J.; Sønderby, C.K.; Sønderby, S.K.; Nielsen, H.; Winther, O.J.B. DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **2017**, *33*, 3387–3395. [[CrossRef](#)]
49. Yu, C.S.; Chen, Y.C.; Lu, C.H.; Hwang, J.K.J.P.S. Prediction of protein subcellular localization. *Proteins* **2006**, *64*, 643–651. [[CrossRef](#)]
50. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
51. Felsenstein, J.J.E. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [[CrossRef](#)] [[PubMed](#)]

52. Howe, K.L.; Contreras-Moreira, B.; De Silva, N.; Maslen, G.; Akanni, W.; Allen, J.; Alvarez-Jarreta, J.; Barba, M.; Bolser, D.M.; Cambell, L. Ensembl genomes 2020—Enabling non-vertebrate genomic research. *Nucleic Acids Res.* **2019**, *48*, D689–D695. [CrossRef]
53. Wei, K.; Pan, S.; Li, Y. Functional characterization of maize C<sub>2</sub>H<sub>2</sub> zinc-finger gene family. *Plant Mol. Biol. Rep.* **2016**, *34*, 761–776. [CrossRef]
54. Wu, C.; Ding, X.; Ding, Z.; Tie, W.; Yan, Y.; Wang, Y.; Yang, H.; Hu, W. The class III peroxidase (POD) gene family in Cassava: Identification, phylogeny, duplication, and expression. *Int. J. Mol. Sci.* **2019**, *20*, 2730. [CrossRef] [PubMed]
55. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
56. Chen, C.; Chen, H.; He, Y.; Xia, R. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv* **2018**, 289660. [CrossRef]
57. Bailey, T.L.; Williams, N.; Misleh, C.; Li, W.W. MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* **2006**, *34*, W369–W373. [CrossRef]
58. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [CrossRef]
59. You, F.M.; Huo, N.; Gu, Y.Q.; Luo, M.-c.; Ma, Y.; Hane, D.; Lazo, G.R.; Dvorak, J.; Anderson, O.D. BatchPrimer3: A high throughput web application for pcr and sequencing primer design. *BMC Bioinform.* **2008**, *9*, 253. [CrossRef]
60. Peden, J.F. Analysis of Codon Usage. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 2000.
61. Zhang, Y.; Ali, U.; Zhang, G.; Yu, L.; Fang, S.; Iqbal, S.; Li, H.; Lu, S.; Guo, L. Transcriptome analysis reveals genes commonly responding to multiple abiotic stresses in rapeseed. *Mol. Breed.* **2019**, *39*, 1–19. [CrossRef]
62. Available online: <https://bigd.big.ac.cn/> (accessed on 18 October 2020).
63. Shinya, T.; Motoyama, N.; Ikeda, A.; Wada, M.; Kamiya, K.; Hayafune, M.; Kaku, H.; Shibuya, N. Functional characterization of CEBiP and CERK1 homologs in Arabidopsis and rice reveals the presence of different chitin receptor systems in plants. *Plant Cell Physiol.* **2012**, *53*, 1696–1706. [CrossRef] [PubMed]
64. Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **2010**, *5*, 725. [CrossRef]
65. Xu, D.; Zhang, Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys. J.* **2011**, *101*, 2525–2534. [CrossRef] [PubMed]
66. Laskowski, R.A.; Chistyakov, V.V.; Thornton, J.M. PDBsum more: New summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.* **2005**, *33*, D266–D268. [CrossRef] [PubMed]
67. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B. Pubchem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* **2019**, *47*, D1102–D1109. [CrossRef]
68. Wu, Q.; Peng, Z.; Zhang, Y.; Yang, J. COACH-D: Improved protein–ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Res.* **2018**, *46*, W438–W442. [CrossRef]
69. Yang, J.; Roy, A.; Zhang, Y. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* **2013**, *29*, 2588–2595. [CrossRef]
70. Roy, A.; Yang, J.; Zhang, Y. COFACTOR: An accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.* **2012**, *40*, W471–W477. [CrossRef]
71. Brylinski, M.; Skolnick, J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 129–134. [CrossRef]
72. Capra, J.A.; Laskowski, R.A.; Thornton, J.M.; Singh, M.; Funkhouser, T.A. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. *PLoS Comput. Biol.* **2009**, *5*, e1000585. [CrossRef]
73. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]
74. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
75. Liu, B.; Li, J.-F.; Ao, Y.; Qu, J.; Li, Z.; Su, J.; Zhang, Y.; Liu, J.; Feng, D.; Qi, K. Lysin motif-containing proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *Plant Cell* **2012**, *24*, 3406–3419. [CrossRef] [PubMed]
76. Tanaka, K.; Nguyen, C.T.; Liang, Y.; Cao, Y.; Stacey, G. Role of LysM receptors in chitin-triggered plant innate immunity. *Plant Signal. Behav.* **2013**, *8*, e22598. [CrossRef]
77. Altenhoff, A.M.; Studer, R.A.; Robinson-Rechavi, M.; Dessimoz, C. Resolving the ortholog conjecture: Orthologs tend to be weakly, but significantly, more similar in function than paralogs. *PLoS Comput. Biol.* **2012**, *8*, e1002514. [CrossRef] [PubMed]
78. Tombuloglu, G.; Tombuloglu, H.; Cevik, E.; Sabit, H. Genome-wide identification of Lysin-Motif Receptor-Like Kinase (*LysM-RLK*) gene family in *Brachypodium distachyon* and docking analysis of chitin/LYK binding. *Physiol. Mol. Plant Pathol.* **2019**, *106*, 217–225. [CrossRef]
79. Zafar, A.; Ahmad, S.; Naseem, I. Insight into the structural stability of coumestrol with human estrogen receptor  $\alpha$  and  $\beta$  subtypes: A combined approach involving docking and molecular dynamics simulation studies. *RSC Adv.* **2015**, *5*, 81295–81312. [CrossRef]

80. Sharp, P.M.; Tuohy, T.M.; Mosurski, K.R. Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res.* **1986**, *14*, 5125–5143. [[CrossRef](#)]
81. Yang, H.; Bayer, P.E.; Tirnaz, S.; Edwards, D.; Batley, J. Genome-wide identification and evolution of receptor-like kinases (RLKs) and receptor like proteins (RLPs) in *Brassica juncea*. *Biology* **2021**, *10*, 17. [[CrossRef](#)]
82. Xu, J.; Wang, G.; Wang, J.; Li, Y.; Tian, L.; Wang, X.; Guo, W. The lysin motif-containing proteins, Lyp1, Lyk7 and LysMe3, play important roles in chitin perception and defense against *Verticillium dahliae* in cotton. *BMC Plant Biol.* **2017**, *17*, 1–18. [[CrossRef](#)]
83. Johnson, D.A.; Thomas, M.A. The monosaccharide transporter gene family in *Arabidopsis* and rice: A history of duplications, adaptive evolution, and functional divergence. *Mol. Biol. Evol.* **2007**, *24*, 2412–2423. [[CrossRef](#)]
84. Bowers, J.E.; Chapman, B.A.; Rong, J.; Paterson, A.H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **2003**, *422*, 433–438. [[CrossRef](#)] [[PubMed](#)]
85. Xu, G.; Guo, C.; Shan, H.; Kong, H. Divergence of duplicate genes in exon–intron structure. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1187–1192. [[CrossRef](#)]
86. Wang, G.-M.; Yin, H.; Qiao, X.; Tan, X.; Gu, C.; Wang, B.-H.; Cheng, R.; Wang, Y.-Z.; Zhang, S.-L. *F-box* genes: Genome-wide expansion, evolution and their contribution to pollen growth in pear (*Pyrus bretschneideri*). *Plant Sci.* **2016**, *253*, 164–175. [[CrossRef](#)]
87. Bai, C.; Sen, P.; Hofmann, K.; Ma, L.; Goebel, M.; Harper, J.W.; Elledge, S.J. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **1996**, *86*, 263–274. [[CrossRef](#)]
88. Dinant, S.; Clark, A.M.; Zhu, Y.; Vilaine, F.; Palauqui, J.-C.; Kusiak, C.; Thompson, G.A. Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms. *Plant Physiol.* **2003**, *131*, 114–128. [[CrossRef](#)] [[PubMed](#)]
89. Lindemose, S.; O’Shea, C.; Jensen, M.; Skriver, K. Structure, function and networks of transcription factors involved in abiotic stress responses. *Int. J. Mol. Sci.* **2013**, *14*, 5842–5878. [[CrossRef](#)]
90. Dong, C.-J.; Shang, Q.-M. Genome-wide characterization of *phenylalanine ammonia-lyase* gene family in watermelon (*Citrullus lanatus*). *Planta* **2013**, *238*, 35–49. [[CrossRef](#)]
91. Cao, J.; Li, M.; Chen, J.; Liu, P.; Li, Z. Effects of MeJA on *Arabidopsis* metabolome under endogenous JA deficiency. *Sci. Rep.* **2016**, *6*, 37674. [[CrossRef](#)]
92. Haasl, R.J.; Payseur, B.A. Microsatellites as targets of natural selection. *Mol. Biol. Evol.* **2012**, *30*, 285–298. [[CrossRef](#)]
93. Qin, Z.; Wang, Y.; Wang, Q.; Li, A.; Hou, F.; Zhang, L. Evolution analysis of simple sequence repeats in plant genome. *PLoS ONE* **2015**, *10*. [[CrossRef](#)] [[PubMed](#)]
94. Wang, Z.; Qiao, Y.; Zhang, J.; Shi, W.; Zhang, J. Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of *Brassica napus* by small RNA and degradome sequencing. *Gene* **2017**, *619*, 61–70. [[CrossRef](#)]
95. Wu, G.; Park, M.Y.; Conway, S.R.; Wang, J.-W.; Weigel, D.; Poethig, R.S. The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* **2009**, *138*, 750–759. [[CrossRef](#)] [[PubMed](#)]
96. Lu, Y.; Feng, Z.; Liu, X.; Bian, L.; Xie, H.; Zhang, C.; Mysore, K.S.; Liang, J. MiR393 and miR390 synergistically regulate lateral root growth in rice under different conditions. *BMC Plant Biol.* **2018**, *18*, 1–12. [[CrossRef](#)] [[PubMed](#)]
97. Soto-Suárez, M.; Baldrich, P.; Weigel, D.; Rubio-Somoza, I.; San Segundo, B. The *Arabidopsis* miR396 mediates pathogen-associated molecular pattern-triggered immune responses against fungal pathogens. *Sci. Rep.* **2017**, *7*, 1–14. [[CrossRef](#)]
98. Zhao, Q.; Nakashima, J.; Chen, F.; Yin, Y.; Fu, C.; Yun, J.; Shao, H.; Wang, X.; Wang, Z.-Y.; Dixon, R.A. Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in *Arabidopsis*. *Plant Cell* **2013**, *25*, 3976–3987. [[CrossRef](#)] [[PubMed](#)]
99. Patel, P.; Yadav, K.; Srivastava, A.K.; Suprasanna, P.; Ganapathi, T.R. Overexpression of native *Musa*-miR397 enhances plant biomass without compromising abiotic stress tolerance in banana. *Sci. Rep.* **2019**, *9*, 1–15. [[CrossRef](#)]
100. Wang, C.Y.; Zhang, S.; Yu, Y.; Luo, Y.C.; Liu, Q.; Ju, C.; Zhang, Y.C.; Qu, L.H.; Lucas, W.J.; Wang, X. MiR397b regulates both lignin content and seed number in *Arabidopsis* via modulating a laccase involved in lignin biosynthesis. *Plant Biotechnol. J.* **2014**, *12*, 1132–1142. [[CrossRef](#)]
101. Zhang, Y.-C.; Yu, Y.; Wang, C.-Y.; Li, Z.-Y.; Liu, Q.; Xu, J.; Liao, J.-Y.; Wang, X.-J.; Qu, L.-H.; Chen, F. Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. *Nat. Biotechnol.* **2013**, *31*, 848–852. [[CrossRef](#)]
102. Liang, Y.; Zhang, Y.; Xu, L.; Zhou, D.; Jin, Z.; Zhou, H.; Lin, S.; Cao, J.; Huang, L. CircRNA expression pattern and ceRNA and miRNA–mRNA networks involved in anther development in the CMS line of *Brassica campestris*. *Int. J. Mol. Sci.* **2019**, *20*, 4808. [[CrossRef](#)]
103. Shen, D.; Suhrkamp, I.; Wang, Y.; Liu, S.; Menkhaus, J.; Verreet, J.A.; Fan, L.; Cai, D. Identification and characterization of microRNAs in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using *Brassica AA* (*Brassica rapa*) and *CC* (*Brassica oleracea*) as reference genomes. *New Phytol.* **2014**, *204*, 577–594. [[CrossRef](#)]
104. Pant, B.D.; Musialak-Lange, M.; Nuc, P.; May, P.; Buhtz, A.; Kehr, J.; Walther, D.; Scheible, W.-R. Identification of nutrient-responsive *Arabidopsis* and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing. *Plant Physiol.* **2009**, *150*, 1541–1555. [[CrossRef](#)]
105. Wang, J.; Jian, H.; Wang, T.; Wei, L.; Li, J.; Li, C.; Liu, L. Identification of microRNAs actively involved in fatty acid biosynthesis in developing *Brassica napus* seeds using high-throughput sequencing. *Front. Plant Sci.* **2016**, *7*, 1570. [[CrossRef](#)]
106. Sun, M.; Qian, X.; Chen, C.; Cheng, S.; Jia, B.; Zhu, Y.; Sun, X. Ectopic expression of *GsSRK* in *Medicago sativa* reveals its involvement in plant architecture and salt stress responses. *Front. Plant Sci.* **2018**, *9*, 226. [[CrossRef](#)] [[PubMed](#)]

107. Ye, Y.; Ding, Y.; Jiang, Q.; Wang, F.; Sun, J.; Zhu, C. The role of receptor-like protein kinases (RLKs) in abiotic stress response in plants. *Plant Cell Rep.* **2017**, *36*, 235–242. [[CrossRef](#)]
108. He, X.; Zhang, J. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* **2005**, *169*, 1157–1164. [[CrossRef](#)] [[PubMed](#)]
109. Teshima, K.M.; Innan, H. Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics* **2008**, *178*, 1385–1398. [[CrossRef](#)] [[PubMed](#)]
110. Chen, J.; Yu, F.; Liu, Y.; Du, C.; Li, X.; Zhu, S.; Wang, X.; Lan, W.; Rodriguez, P.L.; Liu, X. FERONIA interacts with ABI2-type phosphatases to facilitate signaling cross-talk between abscisic acid and RALF peptide in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5519–E5527. [[CrossRef](#)]
111. Kumar, D.; Kumar, R.; Baek, D.; Hyun, T.-K.; Chung, W.S.; Yun, D.-J.; Kim, J.-Y. *Arabidopsis thaliana* RECEPTOR DEAD KINASE1 functions as a positive regulator in plant responses to ABA. *Mol. Plant.* **2017**, *10*, 223–243. [[CrossRef](#)]
112. Sotelo, T.; Lema, M.; Soengas, P.; Cartea, M.; Velasco, P. In vitro activity of glucosinolates and their degradation products against Brassica-pathogenic bacteria and fungi. *Appl. Environ. Microbiol.* **2015**, *81*, 432–440. [[CrossRef](#)]
113. Jun, Z.; Zhang, Z.; Gao, Y.; Zhou, L.; Fang, L.; Chen, X.; Ning, Z.; Chen, T.; Guo, W.; Zhang, T. Overexpression of *GbRLK*, a putative receptor-like kinase gene, improved cotton tolerance to *Verticillium* wilt. *Sci. Rep.* **2015**, *5*, 1–12. [[CrossRef](#)]
114. Van de Wouw, A.P.; Howlett, B.J. Advances in understanding the *Leptosphaeria maculans*—*Brassica* pathosystem and their impact on disease management. *Can. J. Plant Pathol.* **2020**, *42*, 149–163. [[CrossRef](#)]
115. Jamieson, P.A.; Shan, L.; He, P. Plant cell surface molecular cypher: Receptor-like proteins and their roles in immunity and development. *Plant Sci.* **2018**, *274*, 242–251. [[CrossRef](#)] [[PubMed](#)]
116. Buendia, L.; Girardin, A.; Wang, T.; Cottret, L.; Lefebvre, B. LysM Receptor-like kinase and LysM receptor-like protein families: An update on phylogeny and functional characterization. *Front. Plant Sci.* **2018**, *9*, 1531. [[CrossRef](#)]
117. You, Y.; Zheng, Y.; Wang, J.; Chen, G.; Li, S.; Shao, J.; Qi, G.; Xu, F.; Wang, G.; Chen, Z.-H. Molecular evolution and genome-wide analysis of the SBP-box family in cucumber (*Cucumis sativas*). *Plant Growth Regul.* **2021**, *93*, 175–187. [[CrossRef](#)]
118. Yadav, M.K.; Gajbhiye, S. Genome-wide characterization and identification of synonymous codon usage patterns in *Plasmodium knowlesi*. *bioRxiv* **2021**, 425038. [[CrossRef](#)]
119. Sharp, P.M.; Devine, K.M. Codon usage and gene expression level in *Dictyostelium discoideum*: Highly expressed genes do [prefer] optimal codons. *Nucleic Acids Res.* **1989**, *17*, 5029–5040. [[CrossRef](#)] [[PubMed](#)]