



## OPEN ACCESS

## EDITED BY

Setsuko Komatsu,  
Fukui University of Technology, Japan

## REVIEWED BY

Prateek Gupta,  
Hebrew University of Jerusalem, Israel  
Jun Cui,  
Hunan Normal University, China

## \*CORRESPONDENCE

Nazia Rehman,  
naziarehman@parc.gov.pk  
Muhammad Ramzan Khan,  
mrkhan@parc.gov.pk

†These authors share first authorship

## SPECIALTY SECTION

This article was submitted  
to Plant Genomics,  
a section of the journal  
*Frontiers in Genetics*

RECEIVED 29 April 2022

ACCEPTED 08 November 2022

PUBLISHED 05 December 2022

## CITATION

Bashir S, Rehman N, Fakhar Zaman F,  
Naeem MK, Jamal A, Tellier A, Ilyas M,  
Silva Arias GA and Khan MR (2022),  
Genome-wide characterization of the  
*NLR* gene family in tomato (*Solanum  
lycopersicum*) and their relatedness to  
disease resistance.  
*Front. Genet.* 13:931580.  
doi: 10.3389/fgene.2022.931580

## COPYRIGHT

© 2022 Bashir, Rehman, Fakhar Zaman,  
Naeem, Jamal, Tellier, Ilyas, Silva Arias  
and Khan. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Genome-wide characterization of the *NLR* gene family in tomato (*Solanum lycopersicum*) and their relatedness to disease resistance

Sehrish Bashir<sup>1,2†</sup>, Nazia Rehman<sup>1,2\*†</sup>, Fabia Fakhar Zaman<sup>1,2</sup>,  
Muhammad Kashif Naeem<sup>1</sup>, Atif Jamal<sup>3</sup>, Aurélien Tellier<sup>4</sup>,  
Muhammad Ilyas<sup>4</sup>, Gustavo Adolfo Silva Arias<sup>4</sup> and  
Muhammad Ramzan Khan<sup>1,2\*</sup>

<sup>1</sup>National Institute for Genomics and Advanced Biotechnology, National Agricultural Research Centre, Islamabad, Pakistan, <sup>2</sup>PARC Institute for Advanced Studies in Agriculture, NARC, Islamabad, Pakistan, <sup>3</sup>Crop Disease Research Institute, National Agricultural Research Center, Islamabad, Pakistan, <sup>4</sup>Population Genetics, Department of Life Science Systems, School of Life Sciences, Technical University of Munich, Freising, Germany

Nucleotide-binding leucine-rich-repeat receptors (NLR), the largest group of genes associated with plant disease resistance (R), have attracted attention due to their crucial role in protecting plants from pathogens. Genome-wide studies of NLRs have revealed conserved domains in the annotated tomato genome. The 321 *NLR* genes identified in the tomato genome have been randomly mapped to 12 chromosomes. Phylogenetic analysis and classification of NLRs have revealed that 211 genes share full-length domains categorized into three major clades (CNL, TNL, and RNL); the remaining 110 NLRs share partial domains and are classified in CN, TN, and N according to their motifs and gene structures. The cis-regulatory elements of NLRs exhibit the maximum number of these elements and are involved in response to biotic and abiotic stresses, pathogen recognition, and resistance. Analysis of the phylogenetic relationship between tomato NLRs and orthologs in other species has shown conservation among Solanaceae members and variation with *A. thaliana*. Synteny and *Ka/Ks* analyses of *Solanum lycopersicum* and *Solanum tuberosum* orthologs have underscored the importance of *NLR* conservation and diversification from ancestral species millions of years ago. RNA-seq data and qPCR analysis of early and late blight diseases in tomatoes revealed consistent *NLR* expression patterns, including upregulation in infected compared to control plants (with some exceptions), suggesting the role of NLRs as key regulators in early blight resistance. Moreover, the expression levels of NLRs associated with late blight resistance (*Solyc04g007060 [NRC4]* and *Solyc10g008240 [RIB12]*) suggested that they regulate *S. lycopersicum* resistance to *P. infestans*. These findings provide important fundamental knowledge for understanding *NLR* evolution and diversity and will empower the broader characterization of disease resistance genes for pyramiding through speed cloning to develop disease-tolerant varieties.

## KEYWORDS

*NLR* genes, phylogenetic relationship, synteny, *Alternaria solani*, *Phytophthora infestans*, tomato

## Introduction

Solanaceae is a diverse plant family that contains economically important crops including potato (*Solanum tuberosum*), tomato (*Solanum lycopersicum*), tobacco (*Nicotiana attenuata*), and pepper (*Capsicum annuum*) (Särkinen et al., 2013). Tomato is the most valuable cash crop after potato. Its growth and yield production are hindered by various biotic and abiotic factors. Among biotic factors, multiple pathogen invasions, including fungi, oomycetes, bacteria, viruses, and nematodes, cause a 15% harvest reduction, ultimately affecting national economies (Narusaka et al., 2018). However, tomato plants are susceptible to >200 diseases that can affect the plants at any growth stage. Tomatoes are mainly affected by fungal pathogens including *Phytophthora infestans* and *Alternaria solani*, which cause late and early blight disease, respectively (Qian et al., 2017). These fungal pathogens collectively accounted for 49–91% of tomato yield loss in Pakistan (Vleeshouwers et al., 2011).

Plants have evolved elaborate methods for immunity against pathogen invasion while initiating signaling pathways through the defense layer mechanism (Seo et al., 2016). This mechanism is regulated by cell surface barrier and pattern recognition receptors, which provide conserved pathogen-associated molecular patterns (PAMPs) for broad-spectrum resistance. Moreover, a second defensive layer is driven by intracellular immune receptors that induce effector-triggered immunity (ETI). Most intracellular immune receptors in plants belong to the resistance gene (R-gene) family, which plays a key role in shielding against pathogens and programmed cell death (Qian et al., 2017). The R-gene family includes transmembrane proteins such as receptor-like kinases (RLKs) and receptor-like proteins (RLPs) and kinase-like proteins; i.e., the Pto-gene family and nucleotide-binding-leucine-rich repeat receptor (NLR)-encoding gene family (Andolfo et al., 2013).

A large NLR gene family in the tomato genome has been identified and re-annotated using advanced techniques (Andolfo et al., 2014). These NLRs act as signaling networks by inducing pathogenesis-related proteins, producing an apoptotic hypersensitive response and conferring resistance by recognizing pathogen-effector proteins (Seong et al., 2020). The conserved component of the NLRs is NB-ARC, which comprises a nucleotide-binding site (NB), Apaf-1 (apoptotic protease activating factor 1), resistance proteins (R), and CED4 (cell death protein-4) (Meyers et al., 2003). NLRs are also classified based on multi-domain characterization and further diversified by their N- and C-terminal regions. The N-terminus of the Toll-like/interleukin 1 domain is a TIR-type NLR that initiates downstream signaling through

enhanced disease susceptibility 1 (ESD1). Additionally, the coiled-coil domain at the N-terminus causes non-race-specific disease resistance 1 (NDR1) (McHale et al., 2006). A small proportion of NLR proteins show N-terminus resistance to powdery mildew 8 (RPW8) domains (RNL) (Shao et al., 2014). Another variable region, the LRR domain at the C-terminus, makes the full-length gene sequences TNL, CNL, and RNL, while some genes lack and encode the partial domain sub-families CN, TN, and N. Both variable NLR regions exhibit specificity toward pathogen recognition by interacting with specific ligands to introduce resistance and are implicated in the activation of corresponding pathways (Shi et al., 2021). Moreover, NLR required for cell death (NRC) works as a helper to NLRs, which evolve *via* duplication and establish specificities to diverse pathogens together with sensor NLRs (Wu et al., 2017).

To date, more than 300 R-genes have been cloned from a wide range of plant species for resistance. Among these, 80% of genes encode NLR proteins (Kourelis and Van Der Hoorn, 2018). Twenty NLRs against various pathogens have been identified and cloned within the Solanaceae family. Previous genome-wide studies identified NLRs in tomatoes through a conserved domain framework, providing insight into their evolution and diversification in different sub-families (Vossen et al., 2016). Recent improvements in the tomato genome annotation and the use of the RenSeq technique have provided more information to annotate and characterize many genes and also identify novel genes. Therefore, the present study applied a domain-based framework with genome-wide analyses to identify the new NLRs and understand their genetic role in disease resistance. The findings of this study may enhance the understanding of the dynamic role of NLR resistance against invading pathogens based on their evolutionary relationships and analyses of their promoters. Furthermore, we validated the NLR genes by qPCR and MQTLs to inform the development of new varieties through gene expression regulation and speed cloning based on the pyramiding of resistance genes to achieve durable tolerance.

## Materials and methods

### Identification of NLRs orthologs, motifs, and gene structures and phylogenetic analyses

BlastP was performed using the protein sequences of previously reported domains of the NLR gene family as the

query sequences (van der Biezen and Jones, 1998; Meyers et al., 1999; Xiao et al., 2001). To identify the *NLR* orthologs in the tomato genome, Ensembl Plants (<https://plants.ensembl.org/index.html>) (Bolser et al., 2016) and Sol Genomics Network (SGN) (<https://solgenomics.net/>) were used. The sequences were retrieved and confirmed in the ANNA database (<https://biobigdata.nju.edu.cn/ANNA/>) (Liu et al., 2021). After removing duplicate genes, the tomato *NLR* genes were evaluated for sequence-specific domains using the SMART tool (<http://smart.embl-heidelberg.de/>) (Schultz et al., 1998), conserved domain database (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Marchler-Bauer et al., 2002), and Pfam domain (<https://pfam.xfam.org/>) (Finn et al., 2003). After filtering the sequences based on conserved domains, 321 *NLR* genes were identified using a different pipeline from earlier data (Andolfo et al., 2014; Seo et al., 2016; Seong et al., 2020). MEME v.5.4.1 was used to predict the specific conserved motifs in 321 tomato NLRs (Bailey et al., 2009). Gene structure display server (GSDS2.0) (<http://gsds.gao-lab.org/>) was used to draw the gene structures (Hu et al., 2015). To understand the diversity and relationship among NLRs within the tomato genome, the protein sequences were subjected to multiple sequence alignment using CLUSTALW (<https://www.genome.jp/tools-bin/clustalw>) and a maximum likelihood (ML) tree with 1000 bootstrap replications for intact and partial domain protein sequences in MEGAX (<https://www.megasoftware.net/>). The comprehensive demonstration of the motifs and gene structure diversity of the sequences present in the same clades of the phylogenetic tree were visualized using TBtools (Chen et al., 2018). The molecular and structural features of tomato NLRs were calculated using EXPASY ProtParam (<https://web.expasy.org/protparam/>) (Gasteiger et al., 2005). Sub-cellular localization of the identified NLRs was determined using CELLO v.2.5 (<http://cello.life.nctu.edu.tw/>) (Yu et al., 2006).

## Evolutionary relationships between tomato *NLR* genes and other species

BlastP was performed using tomato *NLR* protein sequences on the Ensembl Plants database against *Solanum tuberosum*, *Nicotiana attenuata*, *Capsicum annuum*, *Arabidopsis thaliana*, and the outlier *Saccharomyces cerevisiae* to identify their orthologs. A threshold of identity >70% and E-value <1e<sup>-5</sup> identified 16 orthologs from *A. thaliana*, 34 from *S. tuberosum*, 27 from *C. annuum*, 14 from *N. attenuata*, and two from *S. cerevisiae*. A phylogenetic tree was constructed using MEGAX (<https://www.megasoftware.net/>) through the ML method with 1000 bootstrap replications (Kumar et al., 2018). The final tree annotation was produced using the online iTOL (Interactive Tree of Life) (<https://itol.embl.de/>) (Letunic and Bork, 2016).

## Promoter analysis

Cis-regulatory elements identified in the promoter region of *NLR* genes belonged to nine different sub-classes. Regions 1500 bps upstream of the selected tomato gene IDs were downloaded from the Ensembl Plants database and submitted to the PlantCare online tool (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Lescot et al., 2002). The resultant values for cis-elements were visualized as a heatmap in TBtools.

## Synteny and Ka/Ks analysis

Protein and GFF3 files of tomato were retrieved from the Ensembl Plants database, while potato proteome and gene GFF3 files were obtained from the Sol Genomics Network (SGN) (<https://solgenomics.net/>). BlastP was performed for collinearity prediction. Syntenic gene pairs were identified and visualized as an advanced circos plot in TBtools. The protein and coding sequences of syntenic gene pairs between tomato and potato were acquired in a file for non-synonymous per synonymous (Ka/Ks) substitution rate of genes under selective pressure. The divergence time was calculated for evolutionary origin as  $T = Ks/2r$ , where  $x = 1.5 \times 10^{-8}$  (Wai et al., 2021).

## RNA-seq data analysis

Transcriptomic RNA-seq data of early blight resistance in tomatoes were retrieved from the GEO database under accession number GSE75923 (Sarkar et al., 2017). Raw RNA-seq data on late blight were downloaded from GenBank Sequence Read Archive (SRA, SRP041501) (Zuluaga et al., 2016) to estimate the expression levels. The raw reads were mapped to the reference genome using HISAT2 version 2.1.0. StringTie version 2.1.5 was used to quantify the BAM and GTF files. The Ballgown package in R was used to generate an expression profile in FPKM (fragment per kilobase per million fragments mapped reads). Finally, a heatmap was produced using the FPKM values of *NLR* differentially expressed genes (DEGs) in TBtools.

## Plant materials and fungal disease inoculation

Four tomato (*S. lycopersicum*) genotypes (38037, 38046, 19890, and ROMA) and one wild *S. chilense* genotype (19906) were collected from the BCI (BioResources Conservation Institute), NARC (National Agriculture Research Center), Islamabad. Seeds were first sowed in trays under controlled conditions (glasshouse) at 22–25°C and 90% humidity. After 15 days, the seedlings were transferred to pots. The plants

were inoculated with 30  $\mu$ l of prepared fungal culture suspensions of *Alternaria solani* ( $1 \times 10^3$  per ml) and conidial suspensions of *Phytophthora infestans* ( $5 \times 10^4$  sporangia per ml). The inoculated plants were incubated in a moist growth chamber for 5–7 days at 22–28°C for *A. solani* and 18°C for *P. infestans* with 88–90% relative humidity (Arafa et al., 2017; Akhtar et al., 2019). The disease effects on the infected leaves were estimated as the percentage of affected area (Moghaddam et al., 2019).

## RNA extraction, cDNA synthesis, and quantitative real-time PCR

Leaf samples of control and disease-infected tomato plants were collected 7 days after infection. Total RNAs were extracted using an Invitrogen Pure Link™ RNA Mini Kit according to the manufacturer's instructions. First-strand cDNA libraries were constructed from 0.5  $\mu$ g of total RNA by using a Thermo Scientific Revert Aid Reverse Transcriptase kit (Scientific, 2000). For expression analysis, primers for selected upregulated and downregulated *NLR* genes (RPP13, RPM1, R1B12, R1B16, NRC4, and Rpi-gene) and an internal reference gene elongation factor (EF) were used (Supplementary Table S1). The 10- $\mu$ l real-time PCR reactions (Applied Biosystems StepOnePlus™) included Maxima SYBR Green qPCR master mix, primers, and cDNA (control and infected samples) as a template. The assay was carried out in triplicate and data were analyzed.

## Chromosomal localization and meta-QTL analysis

The co-localization of NLRs with previously reported QTL regions linked to early and late blight resistance was observed using the meta-QTL approach and related QTLs in tomato retrieved from previous studies (Foolad et al., 2002; Zhang et al., 2003; Chaerani et al., 2007; Chen et al., 2014; Haggard et al., 2015; Mazumdar et al., 2021). NLRs and QTL distributions on the tomato chromosomes were visualized using MapChart (Voorrips, 2002). Early- and late blight-resistant QTLs linked to different NLRs were represented in different colors.

## Results

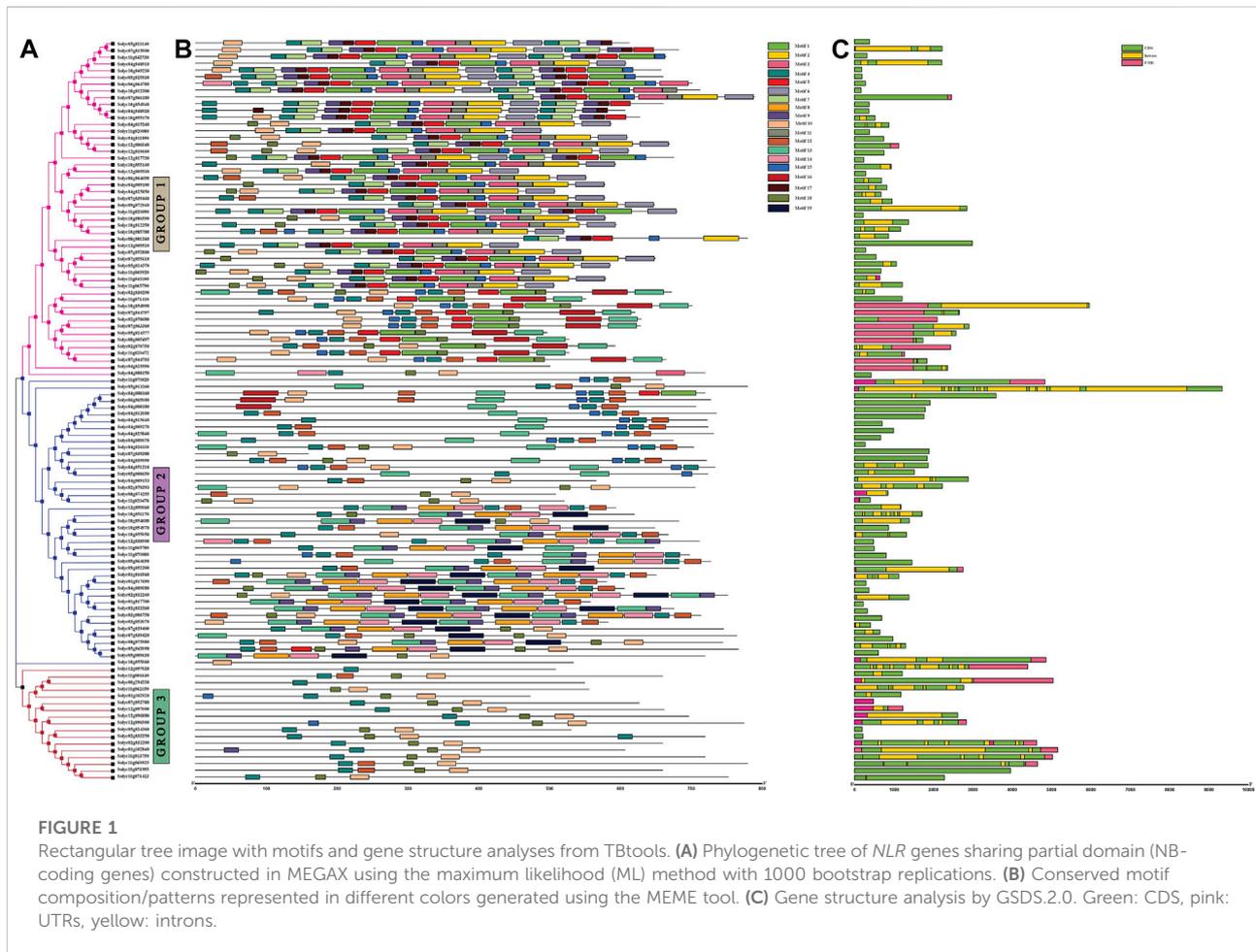
### Identification and evaluation of the *NLR* gene family in tomato

A sequence similarity search was performed in BlastP using *NLR* protein domain sequences (CNL, TNL, and

RNL) as the query. Approximately 455 *NLR* homologs in tomato were identified according to the threshold (>50% identity and E-value  $< 1e^{-5}$ ). The retrieved sequences were subjected to duplicate removal. After filtration, the sequences were submitted to the SMART, Pfam, and CDD databases to confirm the presence of *NLR* domains. A total of 321 *NLR* genes were screened for conserved domains reported in the Pfam database (NB-ARC: PF00931 and LRR: PF00560). This result was consistent with that of previous studies (Andolfo et al., 2014; Andolfo et al., 2021) based on RenSeq. Therefore, the improvement in the present protein domain-based BlastP pipeline is the description of the number of *NLR* genes and their classification according to the variable parts of the domain. Finally, the 321 *NLR* genes were divided into three major subfamilies within the *S. lycopersicum* genome and further categorized according to the domain as intact and partial domain-comprising genes (Supplementary Figure S1). These sequences included 256 CNLs (123 CNL, 35 genes are CNL-type CC-NB, 25 NLcc genes, 57 genes sharing the NB-ARC domain and 16 LRR only), 62 TNLs (21 TNL, 18 genes in TNL-type TIR-NB, and 23 NLtir) and three genes (*Solyc02g090380* (*NRG1*), *Solyc04g079420* (*ADR1*), and *Solyc02g077060*) in the RNL sub-family (Supplementary Table S2). The 321 identified *NLR* genes in tomato with conserved domains were encoding proteins ranging from 64 to 2871 amino acid residues, gene sizes ranging from 194 to 16548 bps, molecular weights ranging from 3330 to 28617 kDa, isoelectric points ranging from 4.5 to 10.27, instability index varying from 10 to 70, and aliphatic indexes varying from 64 to 139 (Supplementary Table S3). Cellular localization determined that the genes were randomly distributed inside the cell independent of chromosome number. For instance, genes present on chromosome 1 were observed in the cytoplasm, plasma membrane, extracellular domain, and nucleus.

### Phylogenetic tree, motif, and gene structure analyses of NB domain-containing NLRs in tomato

A phylogenetic tree of the 110 identified partial-domain NB-coding protein sequences was constructed using MEGAX to determine the phylogenetic relationship of NB domain-containing genes in tomato. The genes were clustered into three groups. The first group represented NB-ARC domain-containing sequences, the second group included CC-NB-containing protein sequences, and the third group showed TIR-NB domain sequences (Figure 1A). The 110 NB-coding protein sequences were then subjected to MEME for analysis of 19 conserved motif patterns (Figure 1B). Major (P-loop, GLPL, and Kinase2) and minor (RNBS-B and RNBS-D) motifs were evaluated in the conserved part of the NB-ARC domain

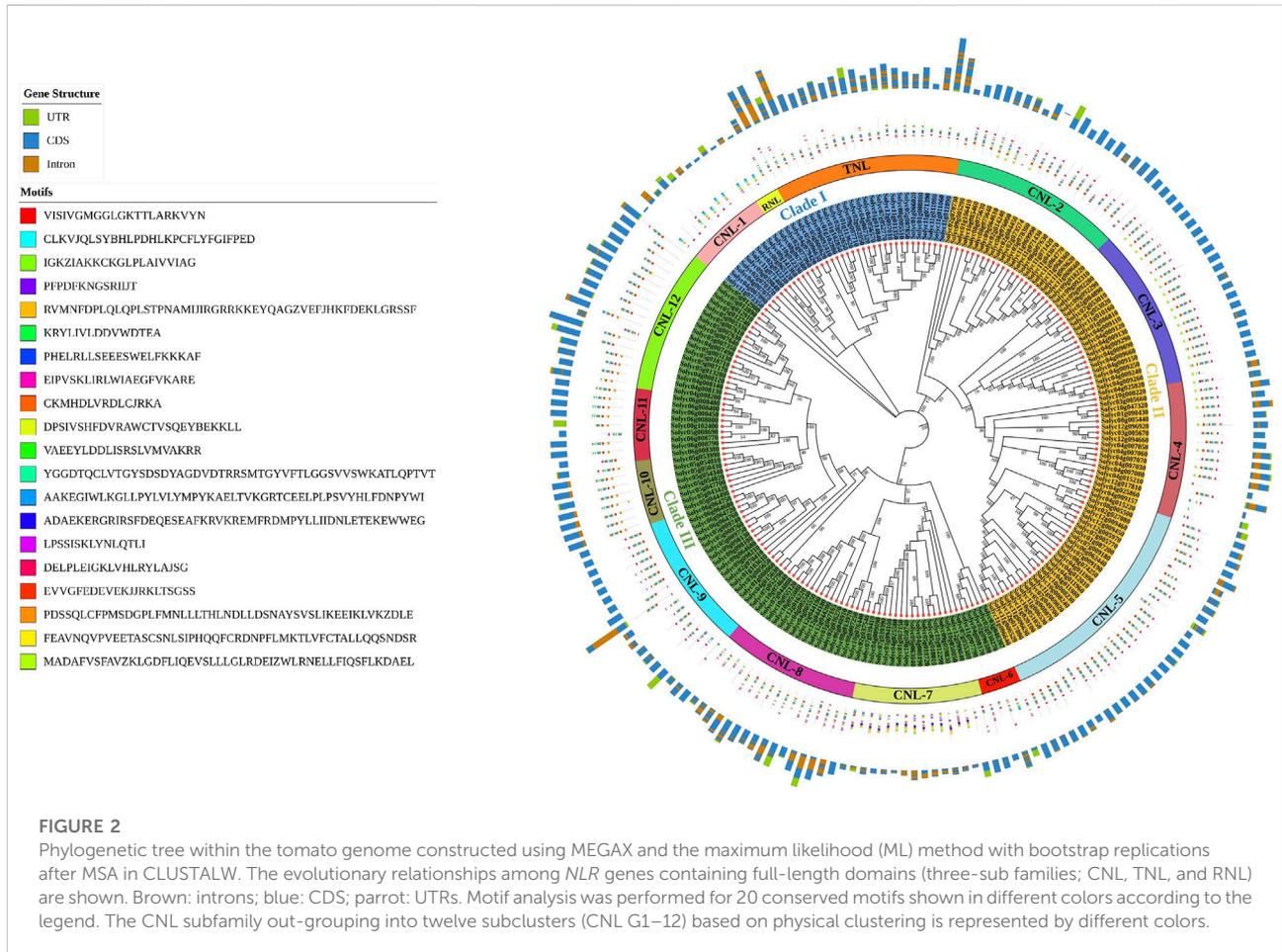


(Supplementary Figure S2). Overall, motifs 13 and 15 were conserved in most of the coding sequences, shown as part of their domains. We observed diversity in the types and numbers of motifs in the genes belonging to different sub-families. The same motif composition was observed in sequences in the same group; for instance, genes with the CC-NB domain showed the same motifs according to their domain profiles. TIR-NB- and NB-domain-containing sequences also represented the conserved motif composition.

The gene structures of the NB-coding sequences were analyzed using GSDS, which determined the number of introns, CDS, and UTRs (Figure 1C). Analysis of gene length variation and structures showed that *Solyc05g013260* contained the maximum of ten introns, while seven introns were present in *Solyc12g097020* in the TN sub-class. Five introns were observed in *Solyc10g054600* and *Solyc12g096900* from the NB and TN subclasses, respectively. Two genes from CN, seven genes from N, and two genes from TN showed at least one intron. Forty-four NB-coding genes had no intron and contained only CDS. The remaining sequences showed one or two UTR regions with one CDS and intron.

## Phylogenetic relationship, motif, and gene structure analyses of *NLR* domain-containing genes in tomato

A total of 211 genes containing full-length CC-NB-LRR, TIR-NB-LRR, and RPW8-NB-LRR domains were subjected to phylogenetic analysis after multiple sequence alignment through CLUSTALW. The results showed that these genes were classified into three major clades and further subdivided into fourteen groups based on sequence similarities with the conserved domain. Clade I contained genes sharing the CNL-1, RNL, and TNL domains. Clade II shared four CNL gene distribution subgroups (CNL-2 to CNL-5). Clade III included seven CNL subgroups (CNL-6 to CNL-12). The CNL sub-class (CNL1-12) out-grouped the numbers of genes present in the other two groups as most of the genes fell into the CNL subfamily and made separate subclasses (Figure 2). Furthermore, the motif patterns were the same for the genes present in each subclass, except for *Solyc11g071390* and *Solyc11g042750*, which had no conserved motifs. Gene structure analysis of the 211 full-length domain sequences showed that 43 tomato genes had only exons



and no introns, whereas *Solyc01g102880* contained the maximum number of introns (17).

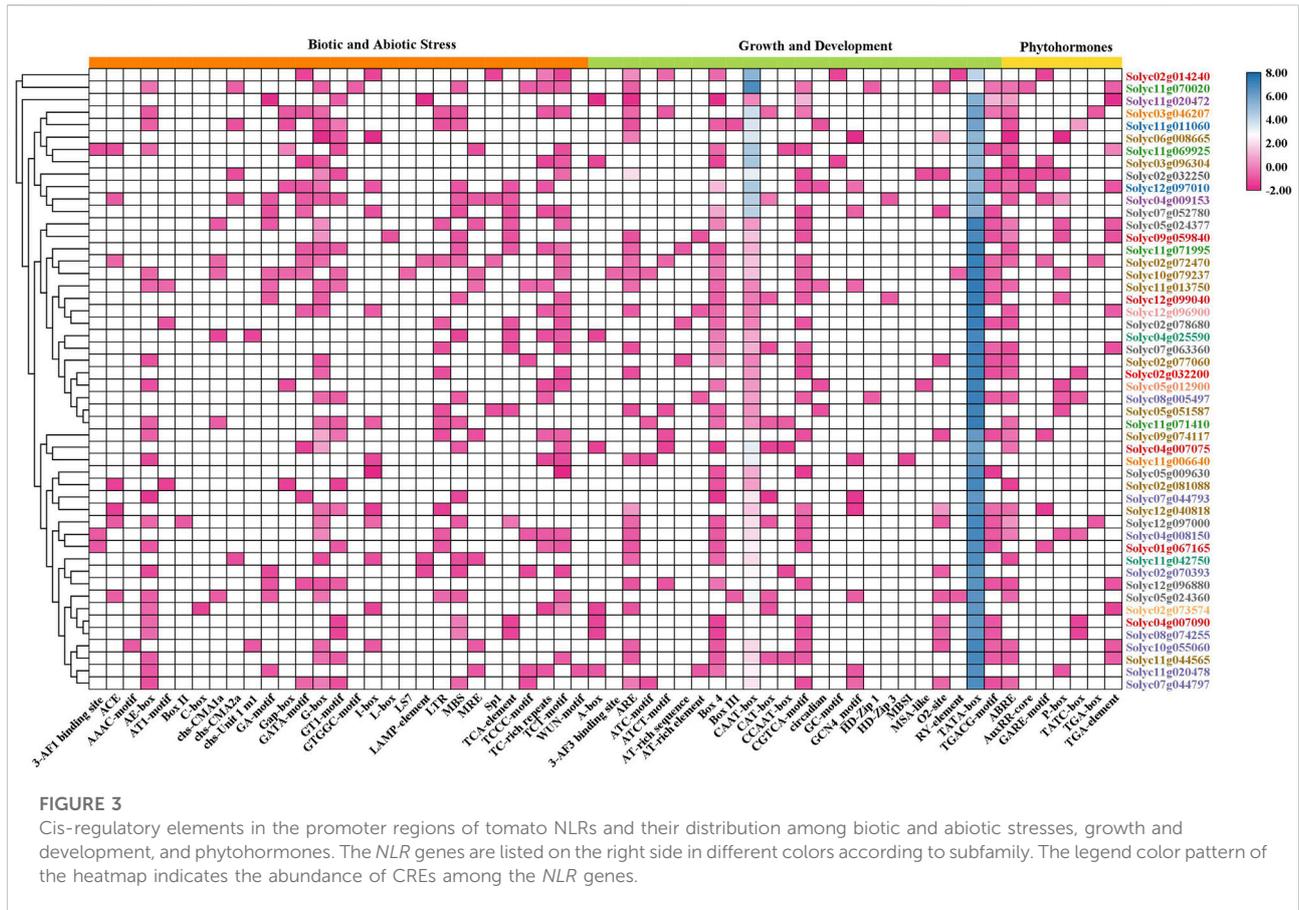
## Analysis of the cis-regulatory elements in the promoter regions of tomato *NLR* genes

Cis-regulatory elements (CRE) present in promoter regions act as transcription start sites and function as regulatory parts of genes. To understand the regulatory role of the identified genes by promoter analysis, 1500 bp regions upstream of the promoter region were retrieved for each representative gene selected from the *NLR* subfamilies. CRE analysis revealed that 29 regulatory elements were involved in biotic and abiotic stresses, 24 in growth and development, and seven in phytohormone response (Figure 3). The heatmap demonstrated that the elements for biotic and abiotic responses, including Box-II, C-box, I-box, LS7, and 3-AF3 binding site, present in each gene's promoter region showed medium expression levels, while the TGA box was associated with phytohormone

response. The maximum abundance of TATA-box and CAAT-box was associated with *NLR* genes for growth and development.

## Phylogenetic analysis of tomato *NLR* genes with other species

To understand the evolutionary relationship of tomato *NLR* genes, one representative tomato gene from each sub-family (TNL, CNL, RNL, CN, N, and TN) was selected. These sequences were subjected to BlastP using threshold values of >70% similarity and E-value <1e<sup>-5</sup> for the identification of orthologs in other related species, including potato, pepper, tobacco, *A. thaliana*, and yeast (selected as an outlier). According to tree topology, three major clades were further subdivided into *NLR* sub-families based on their respective domains, whereas outlier *Saccharomyces cerevisiae* sequences validated the bootstrap replication through the ML method as it appeared as an outer sub-group while the other three major clades presented according to their domain compositions (Figure 4). Clades I and II showed included CNL



subfamily genes (*Solyc11g071995* and *Solyc03g046207*) and their orthologs. These clades are distributed with CNL domain-sharing genes and showed the close relationship of tomato CNL genes with other Solanaceae family members, followed by *A. thaliana*. Clade III showed two sub-clades with major subfamilies distributed with genes from TNL (*Solyc11g013750*) and TN (*Solyc12g097010*) subclasses, which were closely linked to other Solanaceae family members. This clade also contained RNL (*Solyc04g079420*) subfamily genes clustered in association with pepper and potato genes. These results suggested that the NLR gene subfamilies in tomato shared close ancestral relationships with other family members and was associated with the model plant (*Arabidopsis thaliana*). The orthologs in tomato might show the same pathogen interaction as in the Solanaceae family, especially in *Solanum tuberosum* (potato). Furthermore, we also assessed the evolution of the NLR gene family and the number of genes in other species (Supplementary Table S4).

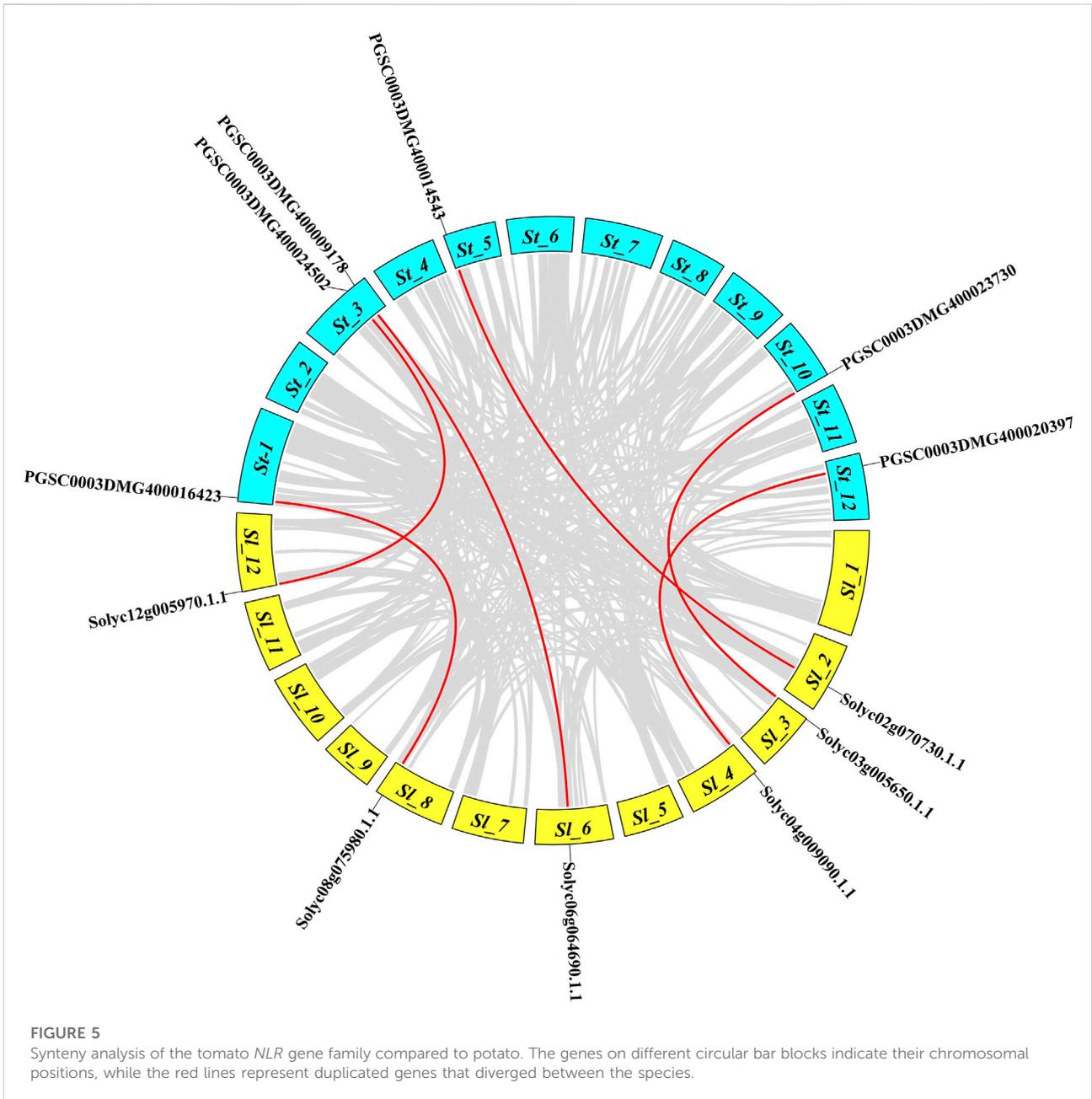
### Synteny and Ka/Ks analysis of NLR genes

The Solanaceae family evolved 52–90 million years ago and includes potato, eggplant, tobacco, tomato, and pepper plants. To

analyze the collinearity relationships among the NLR gene orthologs of tomatoes and potatoes, a microsynteny map was constructed. A total of six orthologs made syntenic gene pairs; for instance, tomato ortholog *Solyc02g070730* present on *Solanum lycopersicum* Sl-2 paired with the potato gene PGSC0003DMG400014543 on *Solanum tuberosum* St-5 (Figure 5). *Solyc12g005970* paired with PGSC0003DMG400024502 on St-3, *Solyc08g075980* paired with PGSC0003DMG400016423 on St-1, *Solyc03g005650* paired with PGSC0003DMG400023730 on St-10, *Solyc04g009090* paired with PGSC0003DMG400020397 on St-12, and *Solyc06g064690* paired with PGSC0003DMG400009178 on St-3. These pairings suggested an evolutionary relationship between the tomato and potato genomes.

To understand the evolutionary origin of genes under selective pressure, synonymous and non-synonymous substitution rates (Ka/Ks) with divergence time (Mya) were calculated for the syntenic gene pairs among tomato and potato (Supplementary Table S5). Ka/Ks values <1 indicate negative/purifying selection, while values of 1 and >1 indicate neutral and positive/natural selection, respectively. All identified gene pairs showed Ka/Ks <1, suggesting that the genes were under negative or purifying selection. With the substitution rate, the divergence time of those genes ranged from 19.5 to 91.6 million years ago (Mya).



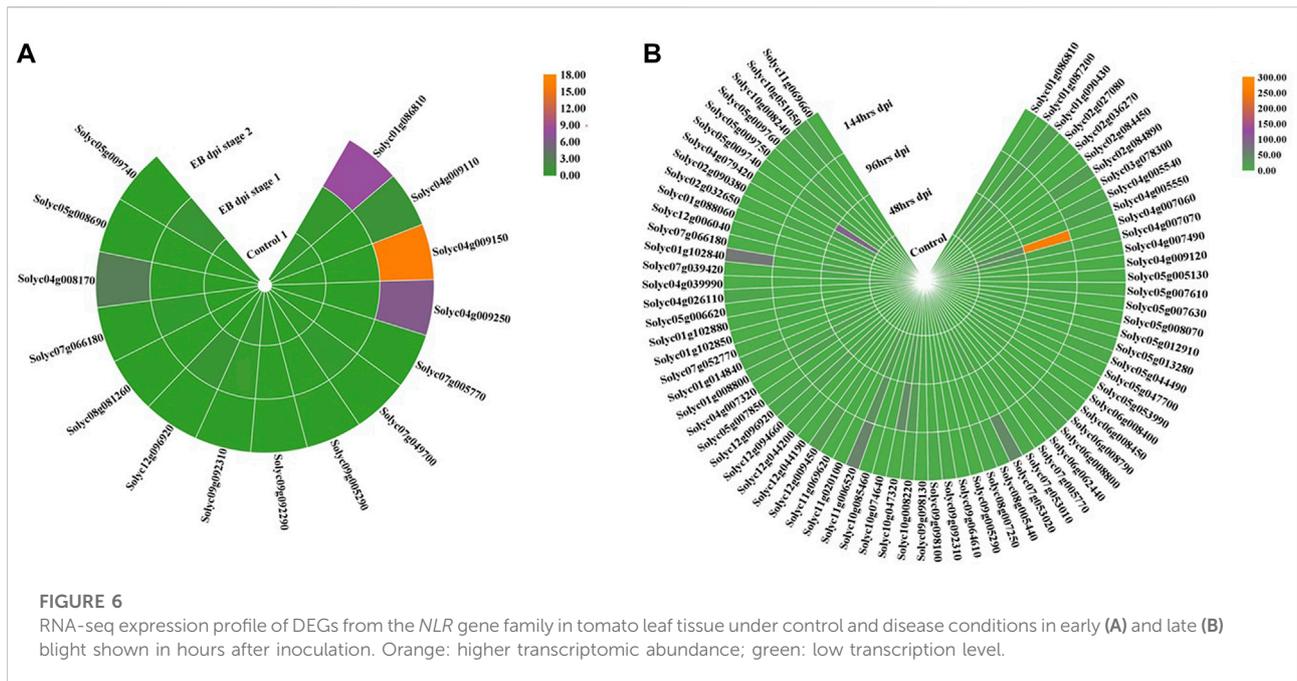


### Expression profiles of tomato NLRs in transcriptomic data

To examine the expression levels of the identified *NLR* genes in tomato during fungal infection, publicly available RNA-seq data of *S. lycopersicum* were retrieved. Data on early blight (EB) and late blight (LB) were obtained from the GEO database (accession number GSE75923) and GenBank Sequence Read Archive (SRA-NCBI) (accession number SRP041501), respectively. Differentially expressed genes (DEGs) were identified against early and late blight diseases in both control

and infected samples (tomato leaf tissues) (Figure 6A,B). Among the 321 *NLRs*, 15 for early blight and 75 for late blight were differentially expressed between control and infected leaves. These findings indicated that *NLRs* play significant roles in conferring tolerance to fungal blight.

*NLR* genes showed both upregulation and downregulation in leaf tissues in the control and infected leaves. In early blight disease, four genes were upregulated in EB-infected tissue compared to the control plants. Two of these genes (*Solyc01g086810* and *Solyc04g009150*) were selected for expression pattern analysis. The remaining 11 DEGs were



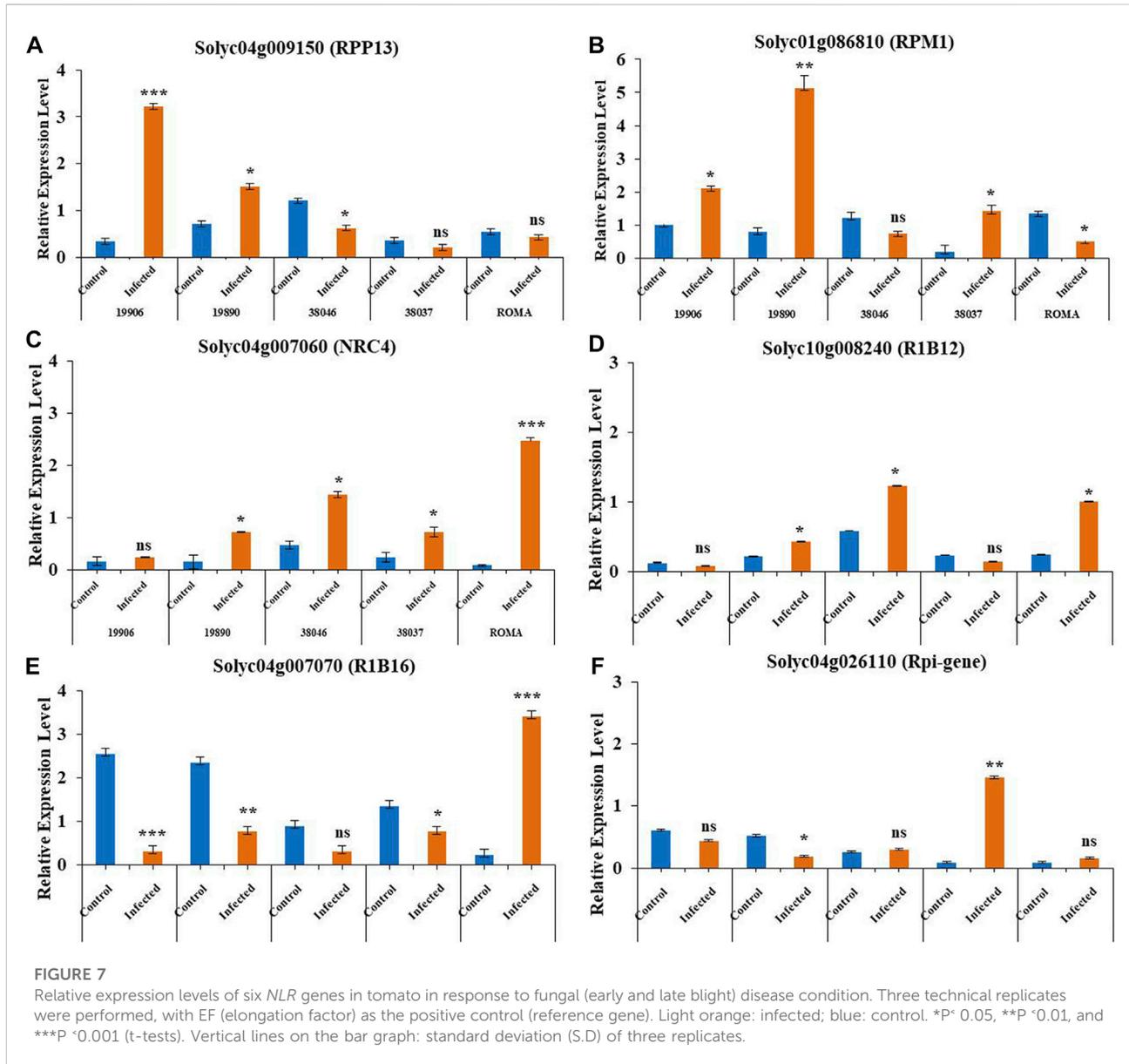
downregulated under disease stress as compared to the controls. In late blight disease, four *NLR* genes showed differential expression (*Solyc04g009110*, *Solyc10g008240*, *Solyc04g007070*, and *Solyc04g007060*). These genes showed upregulation under disease (LB) inoculation in plants. The remaining genes showed lower transcriptional levels in infected leaves than in the control leaves. These findings suggested differential *NLR* behavior in late blight tolerance.

## Validation of expression levels in tomato NLRs infected by different fungal diseases by quantitative PCR

Based on the transcriptomic data, two genes (*Solyc04g009150* [*RPP13*] and *Solyc01g086810* [*RPM1*]) associated with early blight resistance and four genes (*Solyc04g007060* [*NRC4*], *Solyc10g008240* [*RIB12*], *Solyc04g007070* [*RIB16*] and *Solyc04g026110* [*Rpi-gene*]) associated with late blight resistance were selected for qPCR validation. These genes showed dynamic expression patterns in pathogen-infected plants. Between the two early blight resistance NLRs, *Solyc04g009150* (*RPP13*) showed upregulation in infected tolerant wild genotype 19906 (*Solanum chilense*) and tolerant genotype 19890 (*S. lycopersicum*) plants compared to the controls.

However, among moderately susceptible genotypes, 38046 showed increased expression in the control compared to infected plants, while 38037 did not show a significant difference between control and infected plants. The susceptible genotype ROMA also showed no change in

expression pattern between the control and infected samples (Figure 7A). These results suggested similar patterns of expression with the transcriptome levels of these *NLR* genes. In *Solyc01g086810* (*RPM1*) (Figure 7B), infected moderately tolerant genotypes 19890 19906 and moderately susceptible genotype 38037 showed similar expression patterns, in which expression levels were higher in infected plants than in the control. 38046 (moderately susceptible) and ROMA (susceptible) showed downregulation in infected plants. The expression pattern of *Solyc04g007060* (*NRC4*) was also examined in susceptible, moderately susceptible, and tolerant tomato genotypes, in which 19906 (wild genotype) did not show a significant difference between control and infected plants. The remaining genotypes showed higher expression levels in plants with late blight disease than in the control plants (Figure 7C). Similarly, *Solyc10g008240* (*RIB12*) expression was upregulated in response to infection compared to control plants, except for genotypes 19906 and 38037, which showed higher expression levels in control plants, although the difference was not statistically significant (Figure 7D). The inflorescence signals for *Solyc04g007070* (*RIB16*) and *Solyc04g026110* (*Rpi-gene*) were higher in control plants of the examined genotypes than in the infected plants, except for the ROMA and 38037 genotypes (Figures 7E,F). The expression data of these genes revealed patterns consistent with the transcriptome data. The statistical values for relative expression patterns based on  $\Delta\Delta CT$  values in the qPCR validation of genotypes for *NLR* genes in different tomato genotypes are shown in Supplementary Table S7.



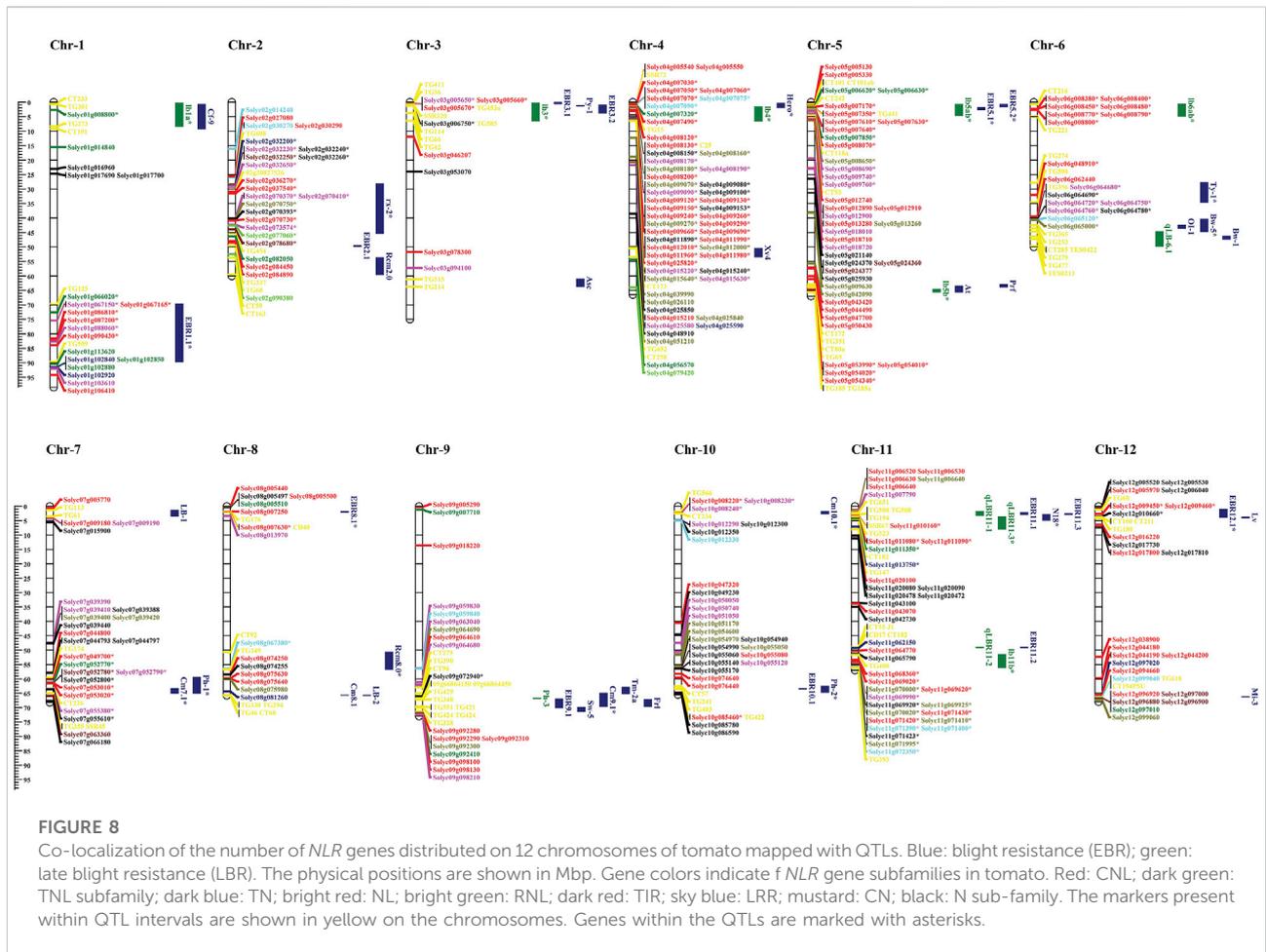
## Genotype-dependent gene expression of tomato NLRs infected by fungal diseases

The expression analysis by qPCR revealed differential expression levels. The wild genotype 19906 showed a strong increase in the expression of early blight resistance NLRs during infection. Furthermore, this genotype did not show a significant difference between control and infected plants and increased expression in the control plants for late blight resistance NLRs. A similar pattern was observed in tolerant genotype 19890. Furthermore, the moderately susceptible varieties 38046 and 38037 showed variable responses with respect to each *NLR* gene, while the susceptible variety ROMA showed significantly increased response for three genes (*Solyc04g007060*, *Solyc10g008240*, and

*Solyc04g007070*) under disease stress compared to the control and no change in expression levels for *Solyc04g026110* and *Solyc04g009150*. *Solyc01g086810* showed significant responses in the control plants compared to the infected plants. These results indicated the specific regulation of tomato *NLR* genes in disease regimes and that their regulation might depend on the genetic architecture of the tomato genotypes.

## Chromosomal localization of tomato *NLR* genes and meta-QTL analysis

The meta-QTL (MQTL) approach is used to identify the quantitative trait loci (QTLs) of complex diseases; i.e., early and



**FIGURE 8**  
Co-localization of the number of *NLR* genes distributed on 12 chromosomes of tomato mapped with QTLs. Blue: blight resistance (EBR); green: late blight resistance (LBR). The physical positions are shown in Mbp. Gene colors indicate *NLR* gene subfamilies in tomato. Red: CNL; dark green: TNL subfamily; dark blue: TN; bright red: NL; bright green: RNL; dark red: TIR; sky blue: LRR; mustard: CN; black: N sub-family. The markers present within QTL intervals are shown in yellow on the chromosomes. Genes within the QTLs are marked with asterisks.

late blight from previous studies. We compared these QTLs to the identified *NLR* genes to reveal the genetic regulators of these QTLs. The numbers of identified *NLR* genes were mapped against reported QTLs for fungal disease resistance and their distribution on 12 tomato chromosomes (Figure 8). The chromosomal localization of the *NLR* genes showed an uneven distribution of genes across the tomato chromosomes. For instance, genes from the CNL subfamily reside on all 12 chromosomes, while TNL genes were absent on chromosomes 3, 6, and 10. The maximum numbers of *NLR* genes were present on chromosomes 4, 5, and 11.

Meta-QTL analysis has shown that chromosome 1 harbors three QTLs, including two related to early blight (blue) and one related to late blight (green). EBR1.1 QTL contained seven genes co-localized from the CNL, TNL, and NL subfamilies. The late blight-related QTL lb1a was linked to one gene (*Solyc1g008800*) from the TNL subfamily.

Chromosome 2 included three QTLs linked to early blight resistance and one QTL linked to late blight resistance. EBR rx-2 QTL shared seventeen genes unevenly from all the categorized sub-classes.

Chromosome 3 included one late blight-related QTL and four early blight-linked QTLs. LBR QTL lb3 co-localized with four genes from the N, CN, and NL sub-families.

Chromosome 4 included two QTLs linked to early blight resistance and one QTL related to late blight resistance. The late blight-related QTL lb4 co-localized with thirty-one *NLR* genes.

Chromosome 5 included six QTLs including two QTLs related to late blight and two related to early blight resistance. QTL Lb5b co-localized with four genes belonging to the CNL subfamily, while QTL lb5ab contained 11 genes from the CNL, TNL, and NL sub-classes. Moreover, EBR QTL EBR5.1 co-localized with five genes from CNL and one TNL subfamily.

Chromosome 6 included six QTLs, two linked to the late blight resistance and four to early blight resistance. LBR QTL lb6ab co-localized with eight *NLR* genes belonging to the CNL and RNL subfamilies.

Chromosome 7 included two early blight-related QTLs and one late blight-related QTL. QTL Ph-1 was linked to late blight resistance in seven co-localized genes while. One early blight-associated QTL (cm7.1) co-localized with two genes from the NL and N sub-classes.

Chromosome 8 included four QTLs linked to early blight resistance and no QTLs associated with late blight. QTL EBR8.1 co-localized with *Solyc08g007630* belonging to the CNL sub-family, while QTL Rcm8.0 shared an interval with *Solyc08g067380* containing the LRR domain.

Chromosome 9 included one late blight-related QTL and four early blight-related QTLs. One early blight-linked QTL, Cm9.1, co-localized with *Solyc09g072940* belonging to the N subfamily.

Chromosome 10 included three MQTLs linked to early blight resistance. QTL Cm10.1 co-localized with three genes from CNL and NL subfamily, while Ph-2 co-localized with *Solyc10g085460* from the CNL subfamily.

Chromosome 11 included eight QTLs, four each linked to early blight and late blight. Among early blight-related QTLs, QTL N18 co-localized with *Solyc11g010160*, which belonged to the CNL subfamily. Among the late blight-related QTLs, qLBR11-3 co-localized with four genes; three from TNL and one from the N subfamily.

Chromosome 12 included three early blight-related QTLs. QTL EBR12.1 co-localized with three *NLR* genes: *Solyc12g009450* and *Solyc12g009460* belonging to the CNL subfamily and *Solyc12g010660* belonging to the N subfamily.

The co-localization of *NLR* genes with the early/late blight disease QTLs on different chromosomes suggests that these *NLRs* may be key regulators of disease resistance among the QTL regions.

## Discussion

Plants are exposed to various biotic and abiotic stresses that affect their growth and yield. Hence, plants have evolved effective immunity against different kinds of pathogens (Gull et al., 2019). Tomato is an economically important fruit crop; however, its growth is dependent on fungicide applications against fungal diseases (*Alternaria solani* and *Phytophthora infestans*), resulting in annual yield loss (Saleem et al., 2015). Resistance genes (R genes) are an important source of signal transduction pathways, which initiate plant–pathogen interactions. The R-gene family includes a large group of *NLR* (nucleotide-binding leucine-rich repeats) genes ubiquitous to plant species that are involved in disease tolerance (Ma et al., 2021). Recently, the evolution and diversification of these *NLR* gene families in numerous plant species have been elucidated based on whole-genome sequencing techniques (Zhang et al., 2016; van Wersch et al., 2020).

The current study identified 321 *NLR*-encoding genes in the tomato genome in modification to previous studies using *NLR* genes from potato and domain-based BlastP framework confirmed by SMART, Interpro, and Pfam approaches. Among these 321 genes, 211 had full-length *NLR* domains and were classified into three subfamilies (CNL, TNL, and RNL), while 110 genes were NB-coding-resistant proteins

containing partial domains TN, CN, and N lacked LRR sequences with improved annotations. Our findings regarding these 321 *NLRs* in tomato supported the classification and distribution of *NLR* genes in the Solanaceae family (Andolfo et al., 2021).

The difference in the number of *NLR* genes between present research and earlier studies within tomato is due to improved genome annotation, selective strategies, and threshold differences for *NLR* gene identification (Seo et al., 2016; Arafa et al., 2017; Shi et al., 2021). Genome size does not correlate with the number of *NLR* genes present in various plants, as potato has a smaller genome than tomato but more *NLRs* due to evolution (Li et al., 2021). Tomato contains fewer *NLRs* than other Solanaceae family members including pepper, potato, eggplant, and tobacco, whereas potato is considered the ancient ancestor of *NLR* genes. A reduced number of *NLR* genes despite the mega genome size can occur due to the loss of a signal transduction component or ecological distribution (Liu et al., 2021).

The TNL and CNL gene ratio in tomato (*S. lycopersicum*) was approximately 1:4, consistent with the predominance of CNLs in other Solanaceae family members. The same ratio was observed in tomato's wild species (*S. pimpinifolium*), with sequence diversity providing broad-spectrum resistance against various pathogens (Qian et al., 2017).

Diverse motif composition was observed with respect to gene subfamily, with genes in the same sub-class, i.e., the CN group, showing the same motif pattern. The identified major (P-loop, GLPL, and Kinase2) and minor (RNBS-B and RNBS-D) motifs in the conserved domain component NB-ARC were consistent with those reported by Seong et al. (2020). The *NLR* genes also showed structural diversity, with CDS, exons, and introns. No introns were present in 40% of the genes. However, alternate exons and exon shuffling may involve versatile protein formation due to the presence of non-coding regions, which also help in exon protection from changes and tend to be conserved for expression (Li et al., 2019).

Phylogenetic analysis based on sequence similarity revealed the classification of *NLR* genes into three subfamilies as reported in previous studies of *NLRs* in *Raphanus sativus* L. (Ma et al., 2021), *S. bicolor* (Zameer et al., 2021), and *D. rotundata* (Zhang et al., 2020). These subfamilies (CNL, TNL, and RNL) were further sub-clustered according to their functionally conserved domains. The CNL subfamily out-grouped into 12 sub-clusters, while TNL and RNL contained two sub-clusters (Andolfo et al., 2014). Two *NLR* genes, ADR1 and NGR1, with the RPW8 domain in their N-terminus region were potentially responsive against powdery mildew, a fungal disease. The RNL subfamily is considered part of the CNL and did not form a separate clade in some plants, whereas the tomato genome contained three RNL domain-encoded genes with the same motif composition (Ma et al., 2021). Tree topology contrary to this showed that the RNL clade was not embedded in CNL

genes but rather presented a separate mini-cluster with a functional domain, indicating that NLRs in plants diverged into three major subclasses after originating from a common ancestor (Shao et al., 2019). With the help of ancestral developments and functional divergence, a new source of resistance can be identified, suggesting that the *NLR* gene present in the same cluster would confer resistance against a specific type of pathogen (Andolfo et al., 2013). The NLRs showed evolutionary relationships among Solanaceae family members, with a high level of conservation among *Solanum* species and a weak association with *A. thaliana*. These findings suggested ancestral duplication or divergence (Wu et al., 2021). An outlier, *S. cerevisiae*, was chosen to assess the reliability of the ML bootstrap replication method, as it out-grouped from the rest of the clustered genes.

Regulatory elements are involved in different responses in the promoter regions of *NLR* genes. Cis elements in *NLR* promoters showed responses to biotic and abiotic stresses compared to growth and development and phytohormones. Box-II, C-box, I-box, LS7, and 3-AF3 binding sites were observed in nearly all promoters of genes against biotic and abiotic stresses. These play key roles in plant protection against pathogens, while the TGA box is present for phytohormones. Moreover, TATA and CAAT boxes were present in maximum abundance for growth and development, consistent with results reported in previous studies (Zameer et al., 2021).

The microsynteny map constructed using *NLR* orthologs in potato and tomato occupying specific chromosomal locations comprising 6% of the tomato and potato genomes, which was consistent with the findings reported by Seong et al. (2020). Ka/Ks analysis and divergence time for syntenic gene pairs between the potato and tomato genomes suggested segmental duplicates that diverged 19–63 million years ago. The evolutionary pattern indicated that these sequence homologs underwent negative or purifying selection, as supported by previous findings (Zhang et al., 2020).

Expression validation of differentially expressed NLRs by qPCR showed strongly increased expression in tolerant and wild genotypes in early blight-resistant NLRs in a study on the induction of defense response upon disease inoculation (Jacob et al., 2013). However, the susceptible genotype ROMA showed higher expression in control than in infected plants. Genotypes with higher expression levels in the control than in infected plants support the upregulation of resistance gene pathways by indicating the downregulation of gene expression after pathogen invasion (Di Donato et al., 2017; Wang et al., 2021). In the case of late blight-resistance NLRs, *Solyc04g007060* was upregulated in infected plants compared to controls, although ROMA also showed significantly increased expression in infected plants compared to the control. In contrast, *Solyc04g007070* showed the same pattern of higher transcriptional levels in the control samples of all genotypes than in infected plants, except ROMA, which suggested the signaling

pathway initiation and modulation of R-genes (Wen et al., 2020). *Solyc04g026110* expression did not differ significantly between control and infected plants except for one moderately susceptible genotype (38037) (Wu et al., 2017). *Solyc10g008240* was upregulated in the tolerant genotypes 19890 and 38046; the other genotypes did not show significant responses. The expression patterns of NLRs in the examined genotypes showed a somewhat consistent pattern with that in previous studies, except for some genotypes with higher expression in the control samples. These findings indicated that genotypes resist the effect of the pathogen by downregulating gene expression through the upregulation of resistance pathways. These include SA signaling and the PR-protein pathway, hypersensitive response, and programmed cell death, at the cost of plant growth and development (MacQueen and Bergelson, 2016). While these responses may be affected by RNA-editing sites or factors where amino acid changes affect the codon, like SNP variations are responsible for the opposite response, which may not allow gene upregulation (Wang et al., 2021). The differential expression of NLRs for disease tolerance in wild and moderately susceptible, and tolerant and susceptible cultivated genotypes, suggests the key role of the overexpression of these antifungal genes and gene pyramiding of NLRs against disease management in managing disease (Kumar et al., 2019).

The distribution of *NLR* genes on tomato chromosomes observed in the present study was consistent with that reported by Shi et al. (2021). The maximum number of genes was distributed on chromosomes 4, 5, and 11. A meta-QTL analysis to obtain insight into the genomic organization of late blight resistance loci based on the closest flanking markers revealed that major QTLs co-localize with mapped *NLR* genes on tomato chromosomes (Khahani et al., 2021). These findings suggested that these loci might be involved in resistance to late blight diseases. The late blight-related resistance QTL lb4 co-localized with the *Solyc04g007060* gene, an expression pattern that was validated by qPCR. This gene determines the induction of defense response upon pathogen invasion, suggesting the role of *NLR* genes in disease tolerance and function as positive regulators of resistance (Danan et al., 2011). The identified QTLs can be further fine-mapped to identify genes within promising regions resistant to diseases (Kumar and Nadarajah, 2020).

NLRs are believed to co-evolve rapidly with pathogens, requiring improved understanding of the molecular mechanisms of disease resistance and the development of resistant cultivars with improved breeding techniques.

## Conclusion

The genome-wide identification of NLRs is a potential approach for resistance gene cloning and functional

characterization. Our results extend the understanding of the abundance and diversity of *NLR* genes in tomato, which may inform gene pyramiding after functional validation. Furthermore, the molecular mechanisms of disease resistance, the development of resistant cultivars, and the over-expression of these NLRs through CRISPR technology may be essential approaches to disease tolerance.

## Data availability statement

Publicly available datasets were analyzed in this study. The names of the repositories and accession number(s) can be found in the article/Supplementary Material. All other data in the study are available on request from the corresponding authors.

## Author contributions

AT, MI, and MRK designed the study. NR, SB, FF, and GS performed tomato plantation, data collection, and molecular analyses. SB and NR wrote the manuscript. MKN assisted with the bioinformatics analyses, figures, MQTL analysis, and writing. NR and MKN supervised the study. AJ helped with the identification and isolation of the fungal cultures and the plant inoculation. MRK provided financial and technical support.

## References

- Akhtar, K. P., Ullah, N., Saleem, M. Y., Iqbal, Q., Asghar, M., and Khan, A. R. (2019). Evaluation of tomato genotypes for early blight disease resistance caused by *Alternaria solani* in Pakistan. *J. Plant Pathol.* 101 (4), 1159–1170. doi:10.1007/s42161-019-00304-8
- Andolfo, G., D'Agostino, N., Frusciantone, L., and Ercolano, M. R. (2021). The tomato interspecific NB-LRR gene arsenal and its impact on breeding strategies. *Genes* 12 (2), 184. doi:10.3390/genes12020184
- Andolfo, G., Jupe, F., Witek, K., Etherington, G. J., Ercolano, M. R., and Jones, J. D. (2014). Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC Plant Biol.* 14 (1), 1–12. doi:10.1186/1471-2229-14-120
- Andolfo, G., Sanseverino, W., Rombauts, S., Van de Peer, Y., Bradeen, J., Carputo, D., et al. (2013). Overview of tomato (*Solanum lycopersicum*) candidate pathogen recognition genes reveals important *Solanum* R locus dynamics. *New Phytol.* 197 (1), 223–237. doi:10.1111/j.1469-8137.2012.04380.x
- Arafa, R. A., Rakha, M. T., Soliman, N. E. K., Moussa, O. M., Kamel, S. M., and Shirasawa, K. (2017). Rapid identification of candidate genes for resistance to tomato late blight disease using next-generation sequencing technologies. *PLoS One* 12 (12), e0189951. doi:10.1371/journal.pone.0189951
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). Meme suite: Tools for motif discovery and searching. *Nucleic Acids Res.* 37 (2), W202–W208. doi:10.1093/nar/gkp335
- Bolser, D., Staines, D. M., Pritchard, E., and Kersey, P. (2016). Ensembl plants: Integrating tools for visualizing, mining, and analyzing plant genomics data. *Methods Mol. Biol.* 1374, 115–140. doi:10.1007/978-1-4939-3167-5\_6
- Chaerani, R., Smulders, M., Van Der Linden, C., Vosman, B., Stam, P., and Voorrips, R. (2007). QTL identification for early blight resistance (*Alternaria solani*) in a *Solanum lycopersicum* x *S. arcanum* cross. *Theor. Appl. Genet.* 114 (3), 439–450. doi:10.1007/s00122-006-0442-8
- Chen, A.-L., Liu, C.-Y., Chen, C.-H., Wang, J.-F., Liao, Y.-C., Chang, C.-H., et al. (2014). Reassessment of QTLs for late blight resistance in the tomato accession L3708 using a restriction site associated DNA (RAD) linkage map and highly aggressive isolates of *Phytophthora infestans*. *PLoS one* 9 (5), e96417. doi:10.1371/journal.pone.0096417
- Chen, C., Chen, H., He, Y., and Xia, R. (2018). TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv* 10 (1101), 289660. doi:10.1016/j.molp.2020.06.009
- Danan, S., Veyrieras, J. B., and Lefebvre, V. (2011). Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. *BMC Plant Biol.* 11 (1), 16–17. doi:10.1186/1471-2229-11-16
- Di Donato, A., Andolfo, G., Ferrarini, A., Delledonne, M., Ercolano, M., and Ferrarini, A. (2017). Investigation of orthologous pathogen recognition gene-rich regions in solanaceous species. *Genome* 60 (10), 850–859. doi:10.1139/gen-2016-0217
- Finn, R., Griffiths Jones, S., and Bateman, A. (2003). Identifying protein domains with the Pfam database. *Curr. Protoc. Bioinforma.* 1 (1)–25. doi:10.1002/0471250953.bi0205s01
- Foolad, M., Zhang, L., Khan, A., Nino-Liu, D., and Lin, G. (2002). Identification of QTLs for early blight (*Alternaria solani*) resistance in tomato using backcross populations of a *Lycopersicon esculentum* x *L. hirsutum* cross. *Theor. Appl. Genet.* 104 (6), 945–958. doi:10.1007/s00122-002-0870-z
- Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., Bairoch, A., et al. (2005). Protein identification and analysis tools on the ExPASy server. *proteomics Protoc. Handb.*, 571–607. doi:10.1385/1-59259-890-0:571
- Gull, A., Lone, A. A., and Wani, N. U. I. (2019). Biotic and abiotic stresses in plants. *Abiotic biotic stress plants*, 1–19. Intechopen, London, UK doi:10.5772/intechopen.77845

## Acknowledgments

We greatly appreciate the *Frontiers* editors and reviewers for handling our manuscript and providing valuable suggestions.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.931580/full#supplementary-material>

- Haggard, J. E., Johnson, E. B., and Clair, D. A. (2015). Multiple QTL for horticultural traits and quantitative resistance to *Phytophthora infestans* linked on *Solanum habrochaites* chromosome 11. *G3-Genes Genomes Genetics* 5 (2), 219–233. doi:10.1534/g3.114.014654
- 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
- Jacob, F., Vernaldi, S., and Maekawa, T. (2013). Evolution and conservation of plant NLR functions. *Front. Immunol.* 4, 297. doi:10.3389/fimmu.2013.00297
- Khahani, B., Tavakol, E., Shariati, V., and Rossini, L. (2021). Meta-QTL and ortho-MQTL analyses identified genomic regions controlling rice yield, yield-related traits and root architecture under water deficit conditions. *Sci. Rep.* 11 (1), 1–18. doi:10.1038/s41598-021-86259-2
- Kourelis, J., and Van Der Hoorn, R. A. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* 30 (2), 285–299. doi:10.1105/tpc.17.00579
- Kumar, A., Jindal, S. K., Dhaliwal, M. S., Sharma, A., Kaur, S., and Jain, S. (2019). Gene pyramiding for elite tomato genotypes against ToLCV (Begomovirus spp.), late blight (*Phytophthora infestans*) and RKN (*Meloidogyne* spp.) for northern India farmers. *Physiol. Mol. Biol. Plants* 25 (5), 1197–1209. doi:10.1007/s12298-019-00700-5
- Kumar, I. S., and Nadarajah, K. (2020). A meta-analysis of quantitative trait loci associated with multiple disease resistance in rice (*Oryza sativa* L.). *Plants* 9 (11), 1491. doi:10.3390/plants9111491
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). Mega X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35 (6), 1547–1549. doi:10.1093/molbev/msy096
- Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30 (1), 325–327. doi:10.1093/nar/30.1.325
- Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44 (1), W242–W245. doi:10.1093/nar/gkw290
- Li, M., Wang, R., Liang, Z., Wu, X., and Wang, J. (2019). Genome-wide identification and analysis of the EIN3/EIL gene family in allotetraploid *Brassica napus* reveal its potential advantages during polyploidization. *BMC Plant Biol.* 19 (1), 110–116. doi:10.1186/s12870-019-1716-z
- Li, Q., Jiang, X.-M., and Shao, Z.-Q. (2021). Genome-wide analysis of NLR disease resistance genes in an updated reference genome of barley. *Front. Genet.* 8, 694682. doi:10.3389/fgene.2021.694682
- Liu, Y., Zeng, Z., Zhang, Y.-M., Li, Q., Jiang, X.-M., Jiang, Z., et al. (2021). An angiosperm NLR Atlas reveals that NLR gene reduction is associated with ecological specialization and signal transduction component deletion. *Mol. Plant* 14 (12), 2015–2031. doi:10.1016/j.molp.2021.08.001
- Ma, Y., Chhapekar, S. S., Lu, L., Oh, S., Singh, S., Kim, C. S., et al. (2021). Genome-wide identification and characterization of NBS-encoding genes in *Raphanus sativus* L. and their roles related to *Fusarium oxysporum* resistance. *BMC Plant Biol.* 21 (1), 47–17. doi:10.1186/s12870-020-02803-8
- MacQueen, A., and Bergelson, J. (2016). Modulation of R-gene expression across environments. *J. Exp. Bot.* 67 (7), 2093–2105. doi:10.1093/jxb/erv530
- Marchler-Bauer, A., Panchenko, A. R., Shoemaker, B. A., Thiessen, P. A., Geer, L. Y., and Bryant, S. H. (2002). Cdd: A database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* 30 (1), 281–283. doi:10.1093/nar/30.1.281
- Mazumdar, P., Singh, P., Kethiravan, D., Ramathani, I., and Ramakrishnan, N. (2021). Late blight in tomato: Insights into the pathogenesis of the aggressive pathogen *Phytophthora infestans* and future research priorities. *Planta* 253 (6), 119–124. doi:10.1007/s00425-021-03636-x
- McHale, L., Tan, X., Koehl, P., and Michelmore, R. W. (2006). Plant NBS-LRR proteins: Adaptable guards. *Genome Biol.* 7 (4), 1–11. doi:10.1186/gb-2006-7-4-212
- Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W., and Young, N. D. (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* 20 (3), 317–332. doi:10.1046/j.1365-313x.1999.t01-1-00606.x
- Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* 15 (4), 809–834. doi:10.1105/tpc.009308
- Moghaddam, G. A., Rezayatmand, Z., Esfahani, M. N., and Khozaei, M. (2019). Genetic defense analysis of tomatoes in response to early blight disease, *Alternaria alternata*. *Plant Physiol. biochem.* 142, 500–509. doi:10.1016/j.plaphy.2019.08.011
- Narusaka, M., Yunokawa, H., and Narusaka, Y. (2018). Efficient identification of NLR by using a genome-wide protein domain and motif survey program, Ex-DOMAIN. *Plant Biotechnol.* 18, 177–180. doi:10.5511/plantbiotechnology.18.0418a
- Qian, L.-H., Zhou, G.-C., Sun, X.-Q., Lei, Z., Zhang, Y.-M., Xue, J.-Y., et al. (2017). Distinct patterns of gene gain and loss: Diverse evolutionary modes of NBS-encoding genes in three Solanaceae crop species. *G3* 7 (5), 1577–1585. doi:10.1534/g3.117.040485
- Saleem, M. Y., Akhtar, K. P., Iqbal, Q., Asghar, M., and Shoaib, M. (2015). Development of high yielding and blight resistant hybrids of tomato. *Pak. J. Agric. Sci.* 52 (2).
- Sarkar, D., Maji, R. K., Dey, S., Sarkar, A., Ghosh, Z., and Kundu, P. (2017). Integrated miRNA and mRNA expression profiling reveals the response regulators of a susceptible tomato cultivar to early blight disease. *DNA Res.* 24 (3), 235–250. doi:10.1093/dnares/dsx003
- Särkinen, T., Bohs, L., Olmstead, R. G., and Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evol. Biol.* 13 (1), 214–215. doi:10.1186/1471-2148-13-214
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A.* 95 (11), 5857–5864. doi:10.1073/pnas.95.11.5857
- Scientific, T. (2000). Thermo scientific Revert Aid H minus Reverse Transcriptase. *Thermo Fish. Sci.* 3, 3–6.
- Seo, E., Kim, S., Yeom, S.-I., and Choi, D. (2016). Genome-wide comparative analyses reveal the dynamic evolution of nucleotide-binding leucine-rich repeat gene family among Solanaceae plants. *Front. Plant Sci.* 7, 1205. doi:10.3389/fpls.2016.01205
- Seong, K., Seo, E., Witek, K., Li, M., and Staskawicz, B. (2020). Evolution of NLR resistance genes with noncanonical N-terminal domains in wild tomato species. *New Phytol.* 227 (5), 1530–1543. doi:10.1111/nph.16628
- Shao, Z. Q., Xue, J. Y., Wang, Q., Wang, B., and Chen, J. Q. (2019). Revisiting the origin of plant NBS-LRR genes. *Trends Plant Sci.* 24 (1), 9–12. doi:10.1016/j.tplants.2018.10.015
- Shao, Z. Q., Zhang, Y. M., Hang, Y. Y., Xue, J. Y., Zhou, G. C., Wu, P., et al. (2014). Long-term evolution of nucleotide-binding site-leucine-rich repeat genes: Understanding gained from and beyond the legume family. *Plant Physiol.* 166 (1), 217–234. doi:10.1104/pp.114.243626
- Shi, J. L., Zai, W. S., Xiong, Z. L., Wan, H. J., and Wu, W. R. (2021). NB-LRR genes: Characteristics in three *Solanum* species and transcriptional response to *Ralstonia solanacearum* in tomato. *Planta* 254 (5), 96–14. doi:10.1007/s00425-021-03745-7
- van der Biezen, E. A., and Jones, J. D. (1998). The NB-arc domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* 8 (7), R226–R227. doi:10.1016/s0960-9822(98)70145-9
- van Wersch, S., Tian, L., Hoy, R., and Li, X. (2020). Plant NLRs: The whistleblowers of plant immunity. *Plant Commun.* 1 (1), 100016. doi:10.1016/j.xplc.2019.100016
- Vleeshouwers, V. G., Raffaele, S., Vossen, J. H., Champouret, N., Oliva, R., Segretin, M. E., et al. (2011). Understanding and exploiting late blight resistance in the age of effectors. *Annu. Rev. Phytopathol.* 49, 507–531. doi:10.1146/annurev-phyto-072910-095326
- Voorrips, R. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93 (1), 77–78. doi:10.1093/jhered/93.1.77
- Vossen, J. H., van Arkel, G., Bergervoet, M., Jo, K.-R., Jacobsen, E., and Visser, R. G. (2016). The *Solanum demissum* R8 late blight resistance gene is an Sw-5 homologue that has been deployed worldwide in late blight resistant varieties. *Theor. Appl. Genet.* 129 (9), 1785–1796. doi:10.1007/s00122-016-2740-0
- Wai, A. H., Cho, L. H., Peng, X., Waseem, M., Lee, D. J., Lee, J. M., et al. (2021). Genome-wide identification and expression profiling of Alba gene family members in response to abiotic stress in tomato (*Solanum lycopersicum* L.). *BMC Plant Biol.* 21 (1), 1–21. doi:10.1186/s12870-021-03310-0
- Wang, W., Ye, J., Yu, C., Xie, Q., Wang, X., Yu, H., et al. (2021). Endogenous RNA editing of a nuclear gene BOSS triggers flowering in tomato. *bioRxiv*. doi:10.1101/2021.06.09.447702
- Wen, B., Zhang, F., Wu, X., and Li, H. (2020). Characterization of the tomato (*Solanum lycopersicum*) pectin methylesterases: Evolution, activity of isoforms and expression during fruit ripening. *Front. Plant Sci.* 11, 238. doi:10.3389/fpls.2020.00238
- Wu, C.-H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., et al. (2017). NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci. U. S. A.* 114 (30), 8113–8118. doi:10.1073/pnas.1702041114

- Wu, J.-Y., Xue, J.-Y., and Van de Peer, Y. (2021). Evolution of *NLR* resistance genes in magnoliids: Dramatic expansions of CNLs and multiple losses of TNLs. *Front. Plant Sci.* 12, 777157. doi:10.3389/fpls.2021.777157
- Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M., et al. (2001). Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* 291 (5501), 118–120. doi:10.1126/science.291.5501.118
- Yu, C. S., Chen, Y. C., Lu, C. H., and Hwang, J. K. (2006). Prediction of protein subcellular localization. *Proteins* 64 (3), 643–651. doi:10.1002/prot.21018
- Zameer, R., Sadaqat, M., Fatima, K., Fiaz, S., Zafar, H., Qayyum, A., et al. (2021). Two-component system genes in *Sorghum bicolor*: Genome-wide identification and expression profiling in response to environmental stresses. *Front. Genet.* 2389, 794305. doi:10.3389/fgene.2021.794305
- Zhang, L., Lin, G., Nino-Liu, D., and Foolad, M. R. (2003). Mapping QTLs conferring early blight (*Alternaria solani*) resistance in a *Lycopersicon esculentum* × *L. hirsutum* cross by selective genotyping. *Mol. Breed.* 12 (1), 3–19. doi:10.1023/A:1025434319940
- Zhang, Y.-M., Chen, M., Sun, L., Wang, Y., Yin, J., Liu, J., et al. (2020). Genome-wide identification and evolutionary analysis of NBS-LRR genes from *Dioscorea rotundata*. *Front. Genet.* 11, 484. doi:10.3389/fgene.2020.00484
- Zhang, Y. M., Shao, Z. Q., Wang, Q., Hang, Y. Y., Xue, J. Y., Wang, B., et al. (2016). Uncovering the dynamic evolution of nucleotide binding site leucine rich repeat (NBS-LRR) genes in Brassicaceae. *J. Integr. Plant Biol.* 58 (2), 165–177. doi:10.1111/jipb.12365
- Zuluaga, A. P., Vega Arreguin, J. C., Fei, Z., Matas, A. J., Patev, S., Fry, W. E., et al. (2016). Analysis of the tomato leaf transcriptome during successive hemibiotrophic stages of a compatible interaction with the oomycete pathogen *Phytophthora infestans*. *Mol. Plant Pathol.* 17 (1), 42–54. doi:10.1111/mpp.12260