

# Genome-Wide Identification and Characterization of NBS-encoding Genes in *Raphanus Sativus* L. And Their Roles Related to *Fusarium Oxysporum* Resistance

**Yinbo Ma**

Chungnam National University College of Agriculture and Life Sciences

**Sushil Satish Chhapekar**

Seoul National University College of Agriculture and Life Sciences

**Lu Lu**

Chungnam National University College of Agriculture and Life Sciences

**Sangheon Oh**

Chungnam National University College of Agriculture and Life Sciences

**Sonam Singh**

Chungnam National University College of Agriculture and Life Sciences

**Chang Soo Kim**

Chungnam National University College of Agriculture and Life Sciences

**Seungho Kim**

Neo Seed Co.

**Gyung Ja Choi**

Korea Research Institute of Chemical Technology

**Yong Pyo Lim**

Chungnam National University College of Agriculture and Life Sciences

**Su Ryun Choi** (✉ [srchoi@cnu.ac.kr](mailto:srchoi@cnu.ac.kr))

Chungnam National University College of Agriculture and Life Sciences <https://orcid.org/0000-0002-1401-9875>

---

## Research article

**Keywords:** *Raphanus sativus* L., NBS-encoding gene, gene duplication, evolution, synteny, *Fusarium oxysporum*

**Posted Date:** August 3rd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-46968/v1>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on January 18th, 2021. See the published version at <https://doi.org/10.1186/s12870-020-02803-8>.

# Abstract

**Background:** The nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes are important for plant development and disease resistance. Although genome-wide studies of NBS-encoding genes have been performed in several species, the evolution, structure, expression, and function of these genes remain unknown in radish (*Raphanus sativus* L.). A recently released draft *R. sativus* L. reference genome has facilitated the genome-wide identification and characterization of NBS-encoding genes in radish.

**Results:** A total of 225 NBS-encoding genes were identified in the radish genome, with 202 mapped onto nine chromosomes and the remaining 23 localized on different scaffolds. According to a gene structure analysis, we identified 99 NBS-LRR-type genes and 126 partial NBS-encoding genes. Additionally, 80 and 19 genes respectively encoded an N-terminal Toll/interleukin-like domain and a coiled-coil domain. Furthermore, 72% of the 202 NBS-encoding genes were grouped in 48 clusters distributed in 24 crucifer blocks on chromosomes. The U block on chromosomes R02, R04, and R08 had the most NBS-encoding genes (48), followed by the R (24), D (23), E (23), and F (17) blocks. These clusters were mostly homogeneous, containing NBS-encoding genes derived from a recent common ancestor. Tandem (15 events) and segmental (20 events) duplications were revealed in the NBS family. Comparative evolutionary analyses of orthologous genes reflected the importance of the NBS-LRR gene family during evolution. Moreover, examinations of cis-elements identified 70 major elements involved in responses to methyl jasmonate, abscisic acid, auxin, and salicylic acid. According to RNA-seq expression analyses, 75 NBS-encoding genes contributed to the resistance of radish to *Fusarium* wilt. A quantitative real-time PCR analysis revealed that *RsTNL03* (*Rs093020*) and *RsTNL09* (*Rs042580*) expression positively regulates radish resistance to *Fusarium oxysporum*, in contrast to the negative regulatory role for *RsTNL06* (*Rs053740*).

**Conclusions:** The NBS-encoding gene structures, tandem and segmental duplications, synteny, and expression profiles in radish were elucidated for the first time and compared with those of other *Brassicaceae* species (*Arabidopsis thaliana*, *Brassica oleracea*, and *Brassica rapa*) to clarify the evolution of the NBS gene family. These results may be useful for functionally characterizing NBS-encoding genes in radish.

## Background

Plants contain numerous resistance (R) genes that are vital for immunity against viral, fungal, and bacterial pathogens [1–3]. Specifically, R gene-mediated disease resistance is one of the most important plant mechanisms related to defense against pathogens [1]. The R genes are grouped in the following five functionally diverse classes based on the presence of specific domains: (1) nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes, including coiled-coil (CC)-NBS-LRR (CNL) and Toll/interleukin 1 receptor (TIR)-NBS-LRR (TNL) genes; (2) receptor-like kinases (RLKs); (3) receptor-like transmembrane proteins; (4) serine–threonine kinases; and (5) atypical R genes [4]. The predominant class of R genes includes genes with NBS and LRR domains [1, 5, 6]. To date, over 300 R genes have been detected and

cloned, of which more than 60% encode NBS and LRR domains [7]. The three NBS-LRR subclasses are TNL, CNL, and resistance to powdery mildew 8 (RPW8)-NBS-LRR (RNL), which are distinguished based on the differences in the N-terminal domains in angiosperms [8]. The TNL and CNL proteins are mainly responsible for the recognition of specific pathogens, whereas RNL proteins participate in downstream defense signal transduction pathways [9]. Therefore, Bonardi et al. described RNL proteins as helper NBS-LRRs [10].

The LRR domain, which is located at the C-terminal of plant NBS-LRR proteins, comprises tandem LRRs involved in the detection of invading pathogens [11]. The NBS domain is a functional ATPase domain, with its nucleotide-binding state believed to regulate R protein activities and function as a molecular switch [12, 13]. Although the TIR and CC domains are implicated in signaling and resistance specificity, their associated pathways differ. The TIR domain is mostly involved in self-association and homotypic interactions with other TIR domains [14, 15], whereas the CC domain may be related to protein-protein interactions and signaling [16]. To protect against diverse and rapidly evolving pathogens, a single plant genome usually encodes hundreds of NBS-LRR genes. The data generated in recent whole-genome sequencing analyses have enabled researchers to comprehensively analyze NBS-LRR genes in economically important plants [17–20].

Radish (*Raphanus sativus* L.), which is one of the more prominent members of the family *Brassicaceae*, is an economically valuable root vegetable crop grown worldwide [21]. Radish quality and yield are influenced by biotic stresses, including fungal and bacterial diseases as well as infestations by insect pests [22]. Specifically, Fusarium wilt caused by the soil-borne fungal pathogen *Fusarium oxysporum*, can severely damage radish and numerous other types of vegetables [23]. Because of its broad host range, *F. oxysporum* can survive at relatively high soil temperatures (> 24 °C) and remain viable even in the absence of any host plant, making it difficult to control [24, 25]. In addition to the TNL-type resistance genes *FOC1* and *FocBo1*, which are responsible for the resistance of *Brassica oleracea* to *F. oxysporum*, many NBS-LRR genes related to Fusarium wilt resistance have been recently identified in diverse plant species [26–28].

Other genes belonging to the NBS-LRR family have been detected in various plants, including *Arabidopsis thaliana* [17, 29], *Brassica rapa* [30], chickpea [31], and *Gossypium* species [32]. Unfortunately, to the best of our knowledge, the potential roles of radish NBS-LRR genes related to disease resistance have not been investigated. The available radish genome sequence is a useful resource for the whole-genome identification of transcription factor families [33, 34]. However, the effects of *F. oxysporum* infections on radish NBS-LRR genes and their families remain unexplored. Therefore, combining bioinformatics and gene expression analyses to systematically study the evolution, expression, and potential functions of NBS-LRR genes may help to improve our understanding of the regulatory networks involved in radish plant growth and in response to *F. oxysporum*.

In this study, a genome-wide analysis of the radish genome identified 225 NBS-encoding genes (99 full NBS-LRR and 126 partial NBS genes) divided into two subclasses (CNL and TNL). These genes were

further characterized regarding chromosomal locations, structures, and duplications. Additionally, with a focus on the NBS-LRR genes, we examined the encoded conserved domains as well as phylogenetic relationships and synteny with genes from *A. thaliana*, *B. rapa*, and *B. oleracea*. Furthermore, the pathogen-induced NBS-LRR gene expression profiles indicated that some R genes are highly expressed in various genetic backgrounds [35]. Our results provide crucial insights into the evolution of this gene family in the radish genome. Moreover, the extensive R gene data presented herein may be useful for accelerating future breeding efforts aimed at improving the disease resistance of radish and other *Brassicaceae* crops.

## Results

### Identification and classification of NBS-encoding genes in *R. sativus*

To comprehensively identify potential NBS-encoding genes in radish, the hidden Markov model (HMM) profile NB-ARC (Pfam: PF00931) from the Pfam database was used to screen the protein sequences encoded in the radish genome [36]. A total of 488 gene candidates with an NBS-LRR domain were identified. These candidate NBS-encoding genes were manually screened and functionally annotated according to the closest *A. thaliana* homolog. Finally, 225 non-redundant NBS-encoding R gene candidates were identified in the Rs1.0 genome (Table 1, Additional file 1: Table S1). The NBS-containing candidate proteins were classified into the TNL or CNL subfamilies based on their HMM profiles and the NCBI Conserved Domain Database, with relationships visualized in a phylogenetic tree (Fig. 1, Table 1). The phylogenetic tree clearly distinguished the TNL and CNL genes in two separate clades (Fig. 1). Gene names were assigned based on their domain type and chromosomal position. The TNL subfamily included 80 genes with full-length domains (TIR, NBS, and LRR) as well as 54 TN (TIR and NBS, but no LRR), 25 NL<sub>TNL</sub> (NBS and LRR, but no TIR), and 15 N<sub>TIR</sub> (no TIR or LRR) genes. The remaining 51 genes belonging to the CNL subfamily included 19 with full-length domains (CC, NBS, and LRR) as well as 10 CN (CC and NBS, but no LRR), 9 RN (RPW8 and NBS, but no LRR), 2 NL<sub>CC</sub> (NBS and LRR, but no CC), and 11 N<sub>CC</sub> (no CC or LRR) genes. Genes encoding RPW8-NBS-LRR proteins were not detected in the radish genome. We also analyzed the well-characterized NBS-encoding genes from the genetically closely related plant species *A. thaliana*, *B. rapa*, and *B. oleracea* (Table 1) using the same method. A total of 164 (*A. thaliana*), 212 (*B. rapa*), and 244 (*B. oleracea*) NBS-encoding genes were identified.

### Genomic distribution among radish chromosomes

Of the 225 NBS-encoding genes, 202 were mapped onto nine radish chromosomes, whereas the other 23 genes were located on different scaffolds (Fig. 2, Additional file 1: Table S1). We determined the distribution of CNL, TNL, and partial NBS-encoding genes on different chromosomes (Additional file 2: Fig. S1). Moreover, the TNL genes were almost uniformly distributed on chromosomes R01 to R09, whereas the CNL genes were not detected on chromosomes R03 and R06. Furthermore, chromosome R09 had the most NBS-encoding genes (41), whereas chromosome R03 had the fewest (7). The ratio of radish

TNL:CNL genes was almost 4:1 (80:19), which was consistent with the corresponding ratios in *A. thaliana* (78:18), *B. rapa* (74:20), and *B. oleracea* (91:20) [17].

<b>Table 1:</b> Summary of NBS gene in <i>R. sativus</i> , <i>A. thaliana</i> , <i>B. rapa</i> and <i>B. oleracea</i>					
Predicted Type	Letter code	<i>R.sativus</i>	<i>A.thaliana</i>	<i>B.rapa</i>	<i>B.oleareae</i>
<b>TNL type</b>					
TIR NB LRR	TNL	80	75	81	93
NB LRR	NL <sub>TNL</sub>	25	4	7	11
TIR NB	TN	54	26	38	50
NB	N <sub>TNL</sub>	15	2	5	23
<b>CNL type</b>					
CC NB LRR	CNL	19	18	19	15
RW8 NB LRR	RNL	0	1	1	1
NB LRR	NL <sub>CC</sub>	2	3	15	3
CC NB	CN	10	11	15	16
RW8 NB	RN	9	4	6	12
NB	N <sub>CC</sub>	11	20	25	20
Total		225	164	212	244

There is no known pattern in the chromosomal distribution of NBS-encoding genes, with most of them detected in clusters. This distribution may facilitate sequence exchanges via recombination mispairing. We identified NBS-encoding gene clusters based on previously established criteria that an NBS-encoding cluster should have two or more genes separated by fewer than 200 kb, with no more than eight non-NBS-encoding genes in between [36]. On radish chromosomes, 146 (72%) NBS genes were mapped in 48 clusters, whereas the remaining 55 genes were detected as singletons (Fig. 2 and Additional file S2). Our analysis revealed that chromosome R09 has the most NBS genes (41; 20.30% of the mapped genes) distributed in eight clusters, in addition to nine singletons. Cluster sizes varied across the genome (2–11 genes). Cluster 44 was the largest, with 11 genes belonging to the TNL subfamily.

To clarify the evolutionary relationships among genes, we analyzed the distribution of NBS-LRR genes among the crucifer blocks in the radish genome. Of the 24 identified blocks (Fig. 2), the U block on chromosomes R02, R04, and R08 was the largest (48 NBS-encoding genes), followed by the R block (24 genes), D block (23 genes), E block (23 genes), and F block (17 genes). The U block may be one of the most important in terms of NBS-encoding genes.

# Gene characteristics and structure

The lengths of the genomic and coding sequences of the 225 NBS-encoding genes as well as the length, molecular weight (MW), and isoelectric point (pI) of the corresponding proteins were comprehensively analyzed (Additional file 1: Table S1). The genomic and coding sequence lengths ranged from 336 bp (RsN02) to 11,267 bp (RsTNL06) and from 1,149 bp (RsN02) to 4,899 bp (RsTNL15), respectively. The protein lengths ranged from 111 amino acids (RsN02) to 1,632 amino acids (RsTNL15). There were also significant variations in the MW and pI, which ranged from 12.72 kDa (RsN11) to 182.40 kDa (RsTNL15) and from 4.77 to 9.60, respectively. Additionally, the average MW of the TNL (122.27 kDa) and CNL (99.84 kDa) proteins were markedly different.

To assess the structural diversity of the *R. sativus* NBS-encoding genes, we compared the number of exons. The full-length CNL and TNL genes in the radish genome had an average of 2.42 and 5.26 exons, respectively (Fig. 3), which is consistent with the analyses of *A. thaliana*, *B. rapa*, and *B. oleracea*, in which the CNL and TNL genes have an average of 2.2, 2.3, and 3.4, and 5.3, 5.2, and 6.4 exons, respectively. Moreover, 47.37% of the CNL genes were encoded by a single exon. These results indicate that the number of introns may have increased and decreased during the structural evolution of the two types of NBS-LRR resistance genes in radish.

## Cis-element analysis

Cis-elements in promoters are usually involved in gene regulation. Thus, to further investigate the potential regulatory networks of the NBS-encoding genes, the cis-elements in the sequences 2 kb upstream of the start codon of 225 NBS-encoding genes were analyzed with PlantCARE (Additional file 1: Table S3). A total of 70 cis-elements were detected, including 10 hormone-responsive elements, 33 light-responsive elements, 4 elements related to abiotic stress, 8 elements associated with tissue-specific expression, and 15 other elements. The cis-element DRE, which is a common cis-acting element in promoter and enhancer regions, was detected in the promoter region of 218 NBS-encoding genes. Additionally, an AT-rich fragment, which is a core promoter element located approximately 30 bp upstream of the transcription start site, was detected in 216 NBS-encoding genes. This sequence was considered to be an essential element in the promoter of the NBS-encoding genes. The promoter region of 159 NBS-encoding genes included the CGTCA-motif, which is a cis-acting regulatory element associated with methyl jasmonate-responsiveness. The abscisic acid-responsive element influencing abscisic acid responsiveness was detected in the promoter region of 165 genes, suggesting that NBS-encoding genes are involved in plant responses to pathogen infections. Moreover, a TCA-element, which is involved in salicylic acid responses, was detected in 97 gene promoters, whereas the TGA-element related to auxin responses was identified in 82 gene promoters. Furthermore, the P-box and GARE-motif associated with gibberellin-responsiveness were present in 55 and 40 NBS-encoding gene promoters, respectively. These results may be relevant for developing a method for identifying candidate genes related to disease resistance.

# Conserved motifs and phylogenetic relationships among TNLs and CNLs

To investigate the TNL and CNL gene and protein structures, we built a phylogenetic tree based on the full-length amino acid sequences encoded by the *R. sativus* TNL and CNL genes. The phylogenetic analysis indicated that the radish NBS-LRRs can be divided into two large groups, namely CNL and TNL (Fig. 4 and Additional file 1: Table S4). The RsTNL group comprised four subgroups. Of the 80 RsTIR proteins, 49, 10, 5, and 16 belonged to subgroups RsTNL-1, RsTNL-2, RsTNL-3, and RsTNL-4, respectively. These results were identical to those of phylogenetic analyses of *R. sativus* and *A. thaliana* (Additional file Fig. 2).

To further elucidate the potential functions and diversification of the TNL and CNL genes in *R. sativus*, 20 encoded conserved motifs were identified and numbered 1–20 based on the MEME program (Additional file 1: Tables S4 and S5). The TIR domain was detected in all RsTNLs. Additionally, the RsTNL-1 subgroup members had most of the motifs. The RsTNL-2 and RsTNL-3 proteins lacked motifs 20 and 11, respectively. The proteins in subgroup RsTNL-4 were missing motifs 11, 12, and 16 (Fig. 3). The RsCNL group members mostly had only six motifs, including the CC domain. Common motif compositions were revealed within subgroups. However, regarding the motif types and numbers, there was considerable diversity among subgroups. This suggests the proteins within subgroups are functionally similar.

## Tandem duplication and synteny analyses of NBS-encoding genes

Whole-genome and tandem duplications are critical events for enhancing genome complexity and evolutionary novelty. In the *R. sativus* genome, 34 of 225 NBS-encoding genes (15.11%) were associated with tandem duplications and were distributed in 15 tandem arrays of 2–5 genes (Additional file 1: Table S6). Our data also revealed variability in the number of duplicated genes per tandem duplication event and an uneven distribution of these duplications on five of nine chromosomes. Genes encoding domains (e.g., CNL, TNL, and TN) were present in the tandem arrays. The 14 tandem duplication events detected for chromosome R09 involved five *RsTN* genes. In contrast, single tandem duplication events occurred on chromosomes R02, R06, R07, R08, and R09, each involving two genes. Chromosome R08 had the most tandem arrays (six tandem groups containing 13 genes), reflecting a hot spot for the distribution of NBS-encoding genes. An analysis of our data according to BLASTP and MCScanX methods identified 20 segmental duplication events involving 32 NBS-encoding genes (Fig. 4 and Additional file 1: Table S7).

The  $K_a/K_s$  values (Additional file 1: Table S8) of the pairs of segmentally duplicated genes were less than 1, indicating these genes evolved under negative selection. These results suggest that both tandem and segmental duplication events were a major driving force for the evolutionary expansion of the NBS-encoding genes in the radish genome.

# Comparative synteny analyses of orthologous pairs of NBS-encoding genes

To further investigate the phylogenetic relationships among the radish NBS-encoding genes, we constructed three synteny maps comparing radish with *A. thaliana*, *B. rapa*, and *B. oleracea*. A total of 209 pairs of NBS-encoding genes (Additional file 1: Table S9) had syntenic relationships between *R. sativus* and *A. thaliana*, *B. rapa*, and *B. oleracea*. Specifically, 39 orthologous gene pairs were detected between *R. sativus* and *A. thaliana*, whereas there were 73 pairs between *R. sativus* and *B. rapa* as well as 97 pairs between *R. sativus* and *B. oleracea* (Fig. 5). Of the 39 pairs between *R. sativus* and *A. thaliana*, all were single-copy genes, except for *RsCN06*, which had two copies. Four partial NBS genes (*RsN<sub>TIR</sub>19*, *RsNL<sub>TIR</sub>07*, *RsTN31*, and *RsTN39*) in the radish genome were detected as homologous to *A. thaliana* TNL genes (*AT1G63730*, *AT5G45210*, and *AT1G63860*). The *A. thaliana* partial NBS gene *AT5G18350* corresponded to the complete NBS gene *RsTNL12*. Of the 73 gene pairs between *R. sativus* and *B. rapa*, 46 radish NBS-LRR genes were present as a single copy, whereas there were two copies of 12 genes and three copies of one gene (*RsTN01*). The *Bra027122* (no TIR domain) and *Bra030779* (no CC domain) *B. rapa* genes corresponded to *RsTNL75* and *RsCNL11*, respectively. Moreover, four *RsTN* genes (*RsTN26*, *34*, *39*, and *44*) and *RsNL01* were syntenic with TNL genes (*Bra001160*, *Bra024652*, *Bra027779*, and *Bra027772*), whereas *RsCN02* was syntenic with a CNL gene (*Bra019063*) in the *B. rapa* genome. Overall, we detected 59 radish NBS-encoding genes syntenic with 64 NBS-encoding genes in the *B. rapa* genome (Additional file 1: Table S9). There were about 97 homologous gene pairs between *R. sativus* and *B. oleracea*, among which 47, 19, and 4 NBS-encoding genes in radish were retained as one, two, and three copies, respectively, in *B. oleracea*, with the remaining genes lacking syntenic relationships. The partial NBS-encoding genes *RsTN11*, *RsTN31*, *RsNL07*, and *RsN19* were syntenic with complete TNL genes (*Bo2g010720*, *Bo8g104700*, *Bo9g061200*, and *Bo9g029350*) in the *B. rapa* genome. Additionally, two *RsCNL* genes (*RsCNL03* and *RsCNL05*) as well as four *RsTNL* genes (*RsTNL09*, *RsTNL14*, *RsTNL62*, and *RsTNL72*) corresponded to partial NBS genes (*Bo6g020950*, *Bo1g048080*, *Bo3g154220*, *Bo6g089230*, *Bo2g126980*, and *Bo3g006960*) in the *B. rapa* genome. We detected a total of 70 radish NBS-encoding genes syntenic with 67 NBS-encoding genes in the *B. oleracea* genome (Additional file 1: Table S9). An analysis of the synteny among the radish, *A. thaliana*, *B. rapa*, and *B. oleracea* genomes revealed the complete NBS-LRR genes were more syntenic than the partial genes.

Finally, a comparative analysis of the orthologous pairs of NBS-LRR genes among four species (*R. sativus*, *A. thaliana*, *B. rapa*, and *B. oleracea*) revealed 22 *R. sativus* NBS-encoding genes with corresponding copies in the *A. thaliana*, *B. rapa*, and *B. oleracea* genomes (Fig. 6 and Additional file 1: Table S10). There was a one-to-one relationship among the NBS-encoding genes between the *R. sativus* and *A. thaliana* genomes. Additionally, half of the radish genes had single-copy syntenic genes in the *B. rapa* and *B. oleracea* genomes. Furthermore, the *B. rapa* and *B. oleracea* genomes had two and three copies of *RsTN38*, respectively, as well as three copies of *RsTN01*. Accordingly, these genes may have been important for the evolution of the NBS-LRR gene family.

The Ka/Ks ratio indicates the selective pressure on genes during evolution. We examined the Ka/Ks ratio for the orthologous gene pairs to determine the evolutionary selection patterns of NBS-LRR genes among *R. sativus*, *A. thaliana*, *B. rapa*, and *B. oleracea*. The Ka/Ks ratio reflects the number of non-synonymous substitutions per non-synonymous site (Ka) and the number of synonymous substitutions per synonymous site (Ks). The ratios for the orthologous gene pairs were estimated for each branch of the phylogenetic tree using the KaKs Calculator. The segmentally and tandemly duplicated NBS-LRR gene pairs as well as all orthologous NBS-LRR gene pairs had a Ka/Ks ratio < 1 (Additional file 1: Table S11), suggesting that the radish NBS-LRR gene family might have experienced strong purifying selection pressure during evolution.

### **Radish NBS-LRR gene expression profiles in response to *Fusarium oxysporum* f. sp. raphanin 59 (FOR59)**

To identify NBS-LRR genes responsive to a *FOR59* infection and determine their spatiotemporal expression patterns, we analyzed the transcriptome data for all NBS-LRR genes. The transcriptome data were generated for seedlings of the 'YR4' and 'YR18' genotypes infected with *FOR59*; the infected samples were collected on days 0, 1, 3, and 6 after the inoculation for the subsequent sequencing on the Illumina platform. Furthermore, we extracted the NBS-LRR genes from the generated RNA-seq data. Transcriptome data were obtained for 171 of 225 NBS-encoding genes based on the conserved domains and gene IDs. The remaining unidentified genes may be unexpressed in response to a *FOR59* infection. Among these 171 NBS-encoding genes, 29 were not differentially expressed between 'YR4' and 'YR18', and were minimally expressed at all time intervals. Thus, these genes were excluded from the gene expression analysis. The fragments per kilobase of exon per million fragments mapped (FPKM) values obtained for the 142 remaining NBS-LRR genes varied from 0 to 278.884. A total of 80 NBS-encoding genes had an FPKM value less than 1 FPKM (low expression), whereas there were 84 genes with more than 1 FPKM value, and less than 10 (high expression). Extremely high expression represented in 13 genes with more than 10, and less than 50; 1 gene with more than 50, less than 100; and 1 gene with more than 100, less than 300. The differentially expressed genes following the *FOR59* inoculation were analyzed with the R package Edge, which revealed 36 genes that were more highly expressed in the resistant 'YR4' line than in the susceptible 'YR18' line at various time-points (Fig. 7). Seven genes (*RsCN10*, *RsN21*, *RsN26*, *RsNL19*, *RsTN04*, *RsTNL15*, and *RsTNL38*) were expressed only in 'YR18', whereas six genes (*RsCN05*, *RsCNL17.3*, *RsN22.1*, *RsTN31*, *RsTN32*, and *RsTN43*) were exclusively expressed in 'YR4'. Most importantly, the expression of five genes (*RsN25*, *RsNL07*, *RsTN18*, *RsTNL09*, and *RsTN17*) gradually increased over time in the 'YR4' plants, but remained relatively stable in the 'YR18' plants, implying they may be crucial for the resistance to *Fusarium* wilt. However, the *RsTNL51* expression level was higher in 'YR18' than in 'YR4', and gradually increased during the infection period (Fig. 7g). These findings suggest a total of 75 NBS-LRR genes contribute to the resistance of radish to *Fusarium* wilt, with six genes (*RsN25*, *RsNL07*, *RsTN18*, *RsTNL09*, *RsTN17*, and *RsTNL51*) potentially crucial for the resistance.

### **Responses to FOR59 infections**

The expression patterns of NBS-LRR-encoding genes in 'YR4' and 'YR18' plants were confirmed by a quantitative real-time (qRT)-PCR assay. Specifically, we validated the expression of only the CNL and TNL genes. Of the 19 *RsCNL* and 80 *RsTNL* genes in the radish genome, approximately 40 genes encoded proteins with amino acid differences between 'YR4' and 'YR18'. The expression patterns of these genes in 'YR4' and 'YR18' at 0, 1, 3, 6, 9, and 12 days after inoculation (DAI) were determined, with the 18S rRNA gene used as an internal control. The *RsTNL03* (*Rs093020*) and *RsTNL09* (*Rs042580*) genes on chromosome R02 were significantly more highly expressed in the resistant 'YR4' plants than in the susceptible 'YR18' plants (Fig. 8), indicating these two genes are related to the Fusarium wilt resistance of radish. Interestingly, both genes were similarly expressed in 'YR4' and 'YR18' on day 0, but the infection rapidly up-regulated the expression levels in 'YR4' plants. The expression levels continued to increase until 9 DAI, further suggesting their potential role in Fusarium wilt resistance. Homologs of *RsTNL09* (*Rs042580*) were identified in *A. thaliana* (*AT4G36150*), *B. rapa* (*Bra010552* and *Bra011666*), and *B. oleracea* (*Bo3g154220* and *Bo7g117810*).

## Discussion

The NBS-LRR genes form a large family of stress resistance genes that are ubiquitous in all plant species. Genome-wide analyses of NBS-LRR gene families have been conducted for numerous species with sequenced genomes [17, 20, 35]. In the current study, we identified candidate NBS-LRR genes and studied their distribution, structure, clustering, duplication, synteny, and conservation. We identified 225 non-redundant NBS-encoding R gene candidates in the radish genome (Table 1, Additional file 1: Table S1) using highly stringent HMMER and Pfam approaches. To confirm the accuracy of our method, a similar approach was used to identify NBS-encoding genes in *B. rapa*, *B. oleracea*, and *A. thaliana*. We identified 164, 212, and 244 NBS-encoding genes in *A. thaliana*, *B. rapa*, and *B. oleracea*, respectively, which is consistent with the results of an earlier investigation by Zhang et al. (2016), in which 165 and 204 NBS-encoding genes were detected in *A. thaliana* and *B. rapa*, respectively. Similar to our study, Yu et al. (2014) also identified 239 NBS-encoding genes in *B. oleracea*. However, the number of NBS-LRR genes identified in *A. thaliana*, *B. oleracea*, and *B. rapa* genomes differed in other studies [17, 18, 30]. These discrepancies are likely due to our stringent HMMER E-value as well as the manual checking for genes with partial NBS domains.

On the basis of a phylogenetic analysis of NBS-encoding candidate genes, 225 NBS-encoding genes were classified into the TNL and CNL subfamilies. The TNL subfamily included 80 TNL, 54 TN, 25 NL<sub>TNL</sub>, and 15 N<sub>TIR</sub> genes, with the remaining 51 genes in the CNL subfamily comprising 19 CNL, 10 C, 9 RN, 2 NL<sub>CC</sub>, and 11 N<sub>CC</sub> genes. Genes encoding RPW8-NBS-LRR proteins were not detected in radish. The CNL:TNL gene ratio was approximately 1:4, which is similar to that in *A. thaliana*, *B. oleracea*, and *B. rapa* [17]. This distinct pattern of TNL gene abundance suggests that the TIR domain is more functionally active than the CC domain in *Brassicaceae* species [37].

The NBS-LRR genes were unevenly distributed across the radish genome, with chromosomes R09 and R03 having the most and fewest genes, respectively. Additionally, 23 of the genes were located on diverse

scaffolds (Fig. 2). Similar to other species, the NBS-LRR genes in radish mainly exist in clusters because of their rapid evolution [38, 39]. About 72% of the radish NBS-LRR genes were detected in 48 clusters, which is higher than the corresponding percentages in *A. thaliana* (61.7%), *B. rapa* (59.4%), and *B. oleracea* (60.3%) [17]. Moreover, 11 genes were included in cluster 44 of the radish genome. Gene families may expand because of polyploidizations, and tandem and local duplications are the most commonly evaluated mechanisms underlying gene family expansions [40]. We identified 66 NBS-encoding genes (29.33%) in radish that underwent tandem and segmental duplications, which is lower than the duplication rates for *A. thaliana* (55.7%), *B. rapa* (47.1%), and *B. oleracea* (43.3%) [17]. The TN genes were predominantly present in tandem arrays, and only three pairs of CN genes were detected in tandem arrays 2 (*RsCn02* and *RsCN03*), 4 (*RsCN07* and *RsCN08*), and 8 (*RsCNL11* and *RsCNL12*). Interestingly, *RsCNL14* and *RsTNL59*, which belong to two distinct subgroups (CNL and TNL), had undergone segmental duplications. These two genes were located in cluster 40 of the radish genome. We speculate that the NBS-LRR genes (especially TNL subfamily genes) in the radish genome had undergone inter- and intraspecific replications. Because radish and *A. thaliana* are *Brassicaceae* species, we investigated the crucifer blocks in the radish genome containing the identified NBS-LRR genes. A total of 45 NBS-LRR genes were located in the U block distributed on different chromosomes (R02, R04, and R08). The genes were also relatively abundant in the R (24 genes), D (23 genes), E (23 genes), and F (17 genes) blocks (Fig. 2). These observations suggest that the *R. sativus* genome structure arose following the rearrangement and divergence from a common ancestor with *A. thaliana* [41, 42]. All tandemly duplicated genes were found in the same clusters (i.e., highly similar), whereas most of the segmentally duplicated genes did not form clusters and were located on different chromosomes. The relatively few tandem and segmental duplications of NBS-encoding genes may help to explain why radish has evolved more slowly than other *Brassicaceae* species [43].

The conserved structural domains of the radish TNL and CNL proteins were examined in this study. The NBS-LRR genes encoding the TIR domain were more common than the genes encoding the CC domain in the analyzed species (*R. sativus*, *A. thaliana*, *B. oleracea*, and *B. rapa*). However, the functions of several partial NBS-encoding genes lacking one or two domains (TIR, CC, and LRR) remain unknown, but their presence in various plant species imply they are important [17]. In the radish genome, the average number of exons was higher for TNL genes than for CNL genes, with half of the CNL genes containing only one exon (Fig. 4). This difference between gene types may be due to the conservative nature of CNL gene replications, which involve many regulatory components. This result is also consistent with those reported for *A. thaliana*, *B. rapa*, and *B. oleracea* [17, 20, 35]. Although the TNL and CNL genes are related to signaling and resistance specificity during pathogen recognition, the genes in these two subfamilies vary regarding sequences and associated signaling pathways and they cluster separately according to phylogenetic analyses [44]. Previous studies proved that the TNL group forms four phylogenetic clades in *B. rapa* [18] and *B. oleracea* [35]. Our motif analysis with the MEME program uncovered diverse motif compositions in the different RsTNL subfamilies. For example, the RsTNL-4 subgroup members had lost motifs 11, 12, and 16, whereas these motifs are present in the RsTNL-1 subgroup members. These differences may contribute to the functional divergence among the TNL proteins in radish.

Radish is an agronomically important root vegetable crop. Its genome, which contains triplicated segments, has intermediate characteristics between the *Brassica* A/C and B genomes, suggesting radish originated from a *Brassica* species [34]. According to a comparison with the *Brassica* A (Br), B (Bn), and C (Bo) genomes, the radish genome has been positioned between the *Brassica* A/C and B genomes [34]. Thus, an analysis of the synteny among the NBS-LRR genes of radish, *A. thaliana*, *B. rapa*, and *B. oleracea* may lead to novel insights into the evolutionary characteristics of RsNBS-encoding genes as well as the phylogenetic relationships with the genes in the other three species. In the current study, 39, 73, and 97 orthologous pairs were respectively identified between radish and *A. thaliana*, *B. rapa*, and *B. oleracea* (Fig. 5), implying that *R. sativus* is more closely related to *B. oleracea* than to *A. thaliana* or *B. rapa*. Moreover, we identified genes in the *A. thaliana*, *B. rapa*, and *B. oleracea* genomes that correspond to 22 *R. sativus* NBS-encoding genes (Fig. 6). Additionally, the orthologous gene pairs may have existed before the ancestral divergence, with important roles related to the evolution of the NBS-LRR gene family [45]. Furthermore, we compared the Ka/Ks values of the orthologous gene pairs between *R. sativus* and *A. thaliana*, *B. rapa*, and *B. oleracea* lineages. For the CNL and TNL gene types, there were no significant differences in the orthologous gene pairs among the three species (Additional file 1: Table S11). We speculate that the NBS-LRR genes in *R. sativus*, *A. thaliana*, *B. rapa*, and *B. oleracea* may have been exposed to diverse selection pressures to develop genetic resistance to the same pathogen. Additionally, some pathogens may be specific to certain *Brassica* species.

In a previous study regarding radish resistance to Fusarium wilt, we isolated 68 NBS-RGAs and 46 SRLK-RGAs from two Fusarium wilt-resistant radish inbred lines, and the genetic diversity was analyzed with RGA-specific primers [46]. We also identified a major QTL (*Fwr1*) containing the *ORF4* gene, encoding a serine/arginine-rich protein kinase that may mediate Fusarium wilt resistance [47]. To further explore the mechanism underlying Fusarium wilt resistance, we profiled the expression of NBS-LRR genes by screening transcriptome data for the resistant ('YR4') and susceptible ('YR18') lines at different time-points during an infection (Fig. 7). A total of 75 NBS-LRR genes were identified as likely involved in the Fusarium wilt resistance of the 'YR4' line, including *RsNL07*, *RsTNL09* (chromosome R02), *RsTN17* and *RsTN18* (chromosome R04), and *RsN25* (scaffold), which exhibited gradually up-regulated expression in infected 'YR4' plants. Accordingly, the genes belonging to the same cluster were similarly expressed. The identified genes and expression profiles based on transcriptome data were verified by a qRT-PCR analysis of the 'YR4' and 'YR18' lines at several time-points after plants were inoculated (Fig. 8). Of the analyzed genes, *RsTNL03* and *RsTNL09* were more highly expressed in 'YR4' (resistant) than in 'YR18' (susceptible) at different time-points. The *RsTNL09* (*Rs042580*) gene is homologous to *AT4G36150* in *A. thaliana*, *Bra010552* and *Bra011666* in *B. rapa*, and *Bo3g154220* and *Bo7g117810* in *B. oleracea*. The *Bra010552* gene encodes a TNL protein, and is one of the candidate genes for clubroot resistance in *B. rapa* [48]. The *RsTNL03* gene, which is a homolog of *AT4G16890* in *A. thaliana*, is likely involved in salicylic acid-dependent early plant defense responses [38–40]. Our RNA-seq and qRT-PCR data also indicated that *RsTNL06* (*Rs053740*) was more highly expressed in 'YR18' than in 'YR4' after the *FOR59* infection, suggesting that *RsTNL06* negatively regulates Fusarium wilt resistance in radish. Interestingly, *RsTNL06* was identified as an ortholog of the *B. oleracea* gene *Bo7g106630*, and may contribute to plant

responses to *Fusarium* species [35]. The potential positive and negative regulators of Fusarium wilt resistance that were identified following the comprehensive analyses of RNA-seq and qRT-PCR data will need to be thoroughly functionally characterized to assess their utility for enhancing the Fusarium wilt resistance of radish.

## Conclusions

A comprehensive analysis of the radish genome revealed 225 NBS-encoding genes, including 99 NBS-LRR genes and 126 partial NBS genes. Additionally, of the 202 NBS-encoding genes mapped onto nine chromosomes, 72% were located in clusters. Another 23 genes were located on different scaffolds. An analysis of syntenic and phylogenetic relationships among the NBS-LRR genes from *A. thaliana*, *B. rapa*, and *B. oleracea* provided valuable insights into the evolutionary characteristics of radish NBS-LRR genes. Moreover, 99 complete NBS-LRR genes were grouped in two main families (TNL and CNL), with the TNL genes further divided into four subgroups with highly similar exon–intron structures and motif compositions. The NBS-LRR genes are important for the resistance of radish plants to various pathogens. In this study, we identified 75 NBS-LRR genes that protect radish plants from Fusarium wilt. We also determined that *RsTNL03* (*Rs093020*) and *RsTNL09* (*Rs042580*) enhance the resistance of radish to *F. oxysporum*, whereas *RsTNL06* (*Rs053740*) has the opposite effect. The phylogenetic and gene expression data presented herein may help to clarify NBS-LRR gene functions.

## Methods

### Identification of NBS-encoding genes

To identify the NBS-encoding genes, we downloaded the whole-genome sequence from the radish database (<http://radish-genome.org/>). We applied hmmsearch of the HMMER (version 3.3) program [49] based on the HMM corresponding to the Pfam NBS (NB-ARC) domain (PF00931) to screen and identify the NBS-encoding genes in the radish genome. We also selected proteins with an E value  $< 1e^{-20}$  for sequence alignments with ClustalW [50]. We constructed the radish-specific NBS HMM with the hmmbuild module of HMMER (version 3.3) to rescan the radish protein database, and proteins with an E-value less than 0.01 were selected for further analyses [17, 51].

The N-terminal of NBS-containing proteins usually include the TIR and CC domains, whereas the C-terminal contains the RPW8 and LRR domains. We used the CLC main workbench (version 7.9) [52] and the Pfam (version 32) program [53] to detect domains in the NBS-containing proteins. These results were confirmed with the NCBI Conserved Domains Tool [54] and MEME (Multiple Expression motifs for Motif Elicitation) [55]. The CC domain in the protein sequences was identified with Paircoil2 [56], with a P score cut-off of 0.025. The NBS-encoding genes in the genomes of *A. thaliana* (Araport11), *B. rapa* (version 1.5), and *B. oleracea* (version 1.1) were also identified using the same method. The *A. thaliana* genome was download from the TAIR database ([www.arabidopsis.org](http://www.arabidopsis.org)), whereas the *B. rapa* and *B. oleracea* genomes were downloaded from the database on the BRAD website (<http://brassicadb.org/brad/>).

## Multiple sequence alignment and phylogenetic analysis of NBS-encoding genes

These analyses confirmed the segregation between the two major NBS-encoding subpopulations (TNL and CNL) in the radish genome and clarified the phylogenetic relationships among the genes on the major branches. Multiple full NBS-containing protein sequences were aligned with the MUSCLE [57] program. The MEGA-X [58] program was used for the phylogenetic analysis, which was completed with the maximum likelihood method based on the Whelan and Goldman model [59]. Finally, the maximum likelihood method involving the LG model with 500 bootstrap replicates was used to construct the phylogenetic tree. Two additional phylogenetic trees were constructed with the same method. One tree was used for analyzing the relationships between the CNL and TNL genes in radish, whereas the other trees were used to verify that the radish CNL and TNL genes are consistent with the corresponding *A. thaliana* genes.

## Chromosomal locations, structures, and duplication events of NBS-encoding genes

The Mapchart (version 2.2) software was used to map all identified NBS-encoding genes onto *R. sativus* chromosomes based on their physical positions indicated in the *R. sativus* genome database. Several genes were organized in diverse NBS-LRR clusters, in which at least two NBS-encoding genes were localized in a 200-kb region and were separated by a maximum of eight non-NBS-LRR genes [36].

The Pepstats program ([http://www.ebi.ac.uk/Tools/seqstats/emboss\\_pepstats](http://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats)) was used to calculate the pI and MW of the NBS-containing proteins. The promoter sequence (2,000 bp upstream of the start codon) of each gene was examined with PlantCARE (<http://bioinformatics.psb.ugcn.be/webtools/plantcare/html/>) to identify cis-elements [60].

The tandem and segmental duplications of the NBS-encoding genes in the radish genome were analyzed with the default parameters of MCScanX [61] and TBtools [62]. Gene duplications were confirmed based on the following two criteria: (a) the length of the shorter aligned sequence covered > 70% of the longer sequence; (b) the similarity of the two aligned sequences was > 70% [63]. The non-synonymous (Ka) and synonymous (Ks) substitutions for each duplication event were calculated with the KaKs Calculator (version 2.0) [64].

The predicted TNL and CNL proteins were analyzed with MEME [65], with default iterative cycles and the maximum number of motifs set to 20. Additionally, TBtools was used to visualize the structures of the TNL and CNL genes according to the genomic and coding sequences.

## Analysis of the orthologous gene pairs between *R. sativus* and three *Brassicaceae* species

Orthologous genes are important for investigating the evolutionary associations among diverse species. In this study, we identified the orthologous gene pairs between the *R. sativus* and *A. thaliana*, *B. rapa*, and *B. oleracea* genomes with the MCScanX program. Specifically, the following parameters were used:  $e = 1e^{-20}$ ,  $u = 1$ , and  $s = 5$  [66]. The orthologous pairs of NBS-encoding genes were extracted and the maps

for analyzing synteny were prepared with TBtools. The 24 genome building blocks from the ancestral karyotype were assigned to the radish genome as previously described [34].

### **Plant growth and *Fusarium oxysporum* inoculation**

The 'YR4' and 'YR18' lines included in this study are highly resistant and susceptible to *F. oxysporum*, respectively. All plants were grown in a culture room at 25 °C with a 16-h photoperiod. *Fusarium oxysporum* f. sp. *raphanin 59* was obtained from the Korea Research Institute of Chemical Technology (Daejeon, Korea). For each plant type, three individuals were collected at 0, 1, 3, 6, 9, and 12 DAI and then pooled for the subsequent RNA isolation and qRT-PCR analysis.

### **Total RNA extraction and gene expression analysis**

Total RNA was isolated from the 'YR4' and 'YR18' plants with the RNeasy Mini Kit (QIAGEN, Germany). The RNA quality and quantity were analyzed by agarose gel electrophoresis and with the Nanodrop 2000 spectrophotometer (Thermo Scientific, USA). The purified total RNA (2 µg) was reverse transcribed to cDNA with the oligo-(dT) primer and TOPscript™ reverse transcriptase (Enzymomics Company, Korea). A qRT-PCR assay was performed with the SYBR Green Supermix and the CFV96™ Real-Time System (Bio-Rad Company, USA) with the following conditions: 95 °C for 3 min; 39 cycles of 95 °C for 15 s and 58 °C for 20 s. Relative gene expression levels were calculated according to the  $2^{-\Delta\Delta C_t}$  method [67].

Previously generated transcriptome data for the *FOR59*-infected 'YR4' and 'YR18' lines at four time-points (0, 1, 3, and 6 DAI) were analyzed (Accession number PRJNA643982). The differentially expressed NBS-LRR genes were analyzed with the R package Edge. Additionally, heat maps were produced with TBtools.

## **Abbreviations**

CC

coiled-coil; CNL:coiled-coil nucleotide-binding site–leucine-rich repeat; *FOR*:*Fusarium oxysporum* f. sp. *raphanin 59*; FPKM:fragments per kilobase of exon per million fragments mapped; HMM:hidden Markov model; NBS-LRR:nucleotide-binding site–leucine-rich repeat; RPW8:resistance to powdery mildew 8; TNL:Toll/interleukin 1 receptor nucleotide-binding site–leucine-rich repeat

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

## Availability of data and materials

The datasets generated and analysed during the current study are available in the NCBI repository, <https://www.ncbi.nlm.nih.gov/>, under accessions PRJNA643982 or are available from the corresponding author on request.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry (IPET) through Golden Seed Project (213006-05-4-SB020, 213006-05-4-SB110), funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Services (KFS).

## Author's contributions

SRC, YPL and SSC conceived and designed the study. YM carried out experiments, generated data and wrote draft manuscript. SSC executed data analysis and wrote draft of the manuscript. LL, SOH, SS were participated in phenotype evaluations and performing the experiments. SS, CSK contributed for discussion of bioinformatic analysis. SK provided the plant materials and supported the plant's main characteristics. GJC helped phenotypic investigation of plants against pathotype, *Fusarium oxysporum* f. sp. *raphanin* 59. SRC and YPL participated as a director, supervised this study, data interpretation and modified the manuscript. All authors have read and approved the final manuscript.

## Acknowledgements

We thank William Yajima for editing the English text of a draft of this manuscript. This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry (IPET) through Golden Seed Project (213006-05-4-SB020, 213006-05-4-SB110), funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Services (KFS).

## References

1. Dangl JL, Jones JD: Plant pathogens and integrated defence responses to infection. *Nature*. 2001;411(6839):826-833.
2. Maekawa T, Kufer TA, Schulze-Lefert P: NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology*. 2011;12(9):817.
3. Dangl JL, Horvath DM, Staskawicz BJ: Pivoting the plant immune system from dissection to deployment. *Science*. 2013;341(6147):746-751.

4. Zhong Y, Cheng Z-MM: A unique RPW8-encoding class of genes that originated in early land plants and evolved through domain fission, fusion, and duplication. *Scientific Reports*. 2016;6:32923.
5. Yue JX, Meyers BC, Chen JQ, Tian D, Yang S: Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *New Phytologist*. 2012;193(4):1049-1063.
6. Marone D, Russo MA, Laidò G, De Leonardis AM, Mastrangelo AM: Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. *International Journal of Molecular Sciences*. 2013;14(4):7302-7326.
7. Kourelis J, van der Hoorn RA: Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *The Plant Cell*. 2018;30(2):285-299.
8. Shao Z-Q, Xue J-Y, Wu P, Zhang Y-M, Wu Y, Hang Y-Y, Wang B, Chen J-Q: Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiology*. 2016;170(4):2095-2109.
9. Collier SM, Hamel L-P, Moffett P: Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Molecular Plant-microbe Interactions*. 2011;24(8):918-931.
10. Bonardi V, Tang S, Stallmann A, Roberts M, Cherkis K, Dangl JL: Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proceedings of the National Academy of Sciences*. 2011;108(39):16463-16468.
11. DeYoung BJ, Innes RW: Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*. 2006;7(12):1243-1249.
12. Andersen EJ, Ali S, Byamukama E, Yen Y, Nepal MP: Disease resistance mechanisms in plants. *Genes*. 2018;9(7):339.
13. Van Ooijen G, Mayr G, Kasiem MM, Albrecht M, Cornelissen BJ, Takken FL: Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*. 2008;59(6):1383-1397.
14. Bernoux M, Ve T, Williams S, Warren C, Hatters D, Valkov E, Zhang X, Ellis JG, Kobe B, Dodds PN: Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host & Microbe*. 2011;9(3):200-211.
15. Ve T, Williams SJ, Kobe B: Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis*. 2015;20(2):250-261.
16. van Ooijen G, van den Burg HA, Cornelissen BJ, Takken FL: Structure and function of resistance proteins in solanaceous plants. *Annu Rev Phytopathol*. 2007;45:43-72.
17. Yu J, Tehrim S, Zhang F, Tong C, Huang J, Cheng X, Dong C, Zhou Y, Qin R, Hua W: Genome-wide comparative analysis of NBS-encoding genes between Brassica species and *Arabidopsis thaliana*. *BMC Genomics*. 2014;15(1):3.
18. Lv S, Changwei Z, Tang J, Li Y, Wang Z, Jiang D, Hou X: Genome-wide analysis and identification of TIR-NBS-LRR genes in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) reveal expression patterns to TuMV infection. *Physiological and Molecular Plant Pathology*. 2015;90:89-97.

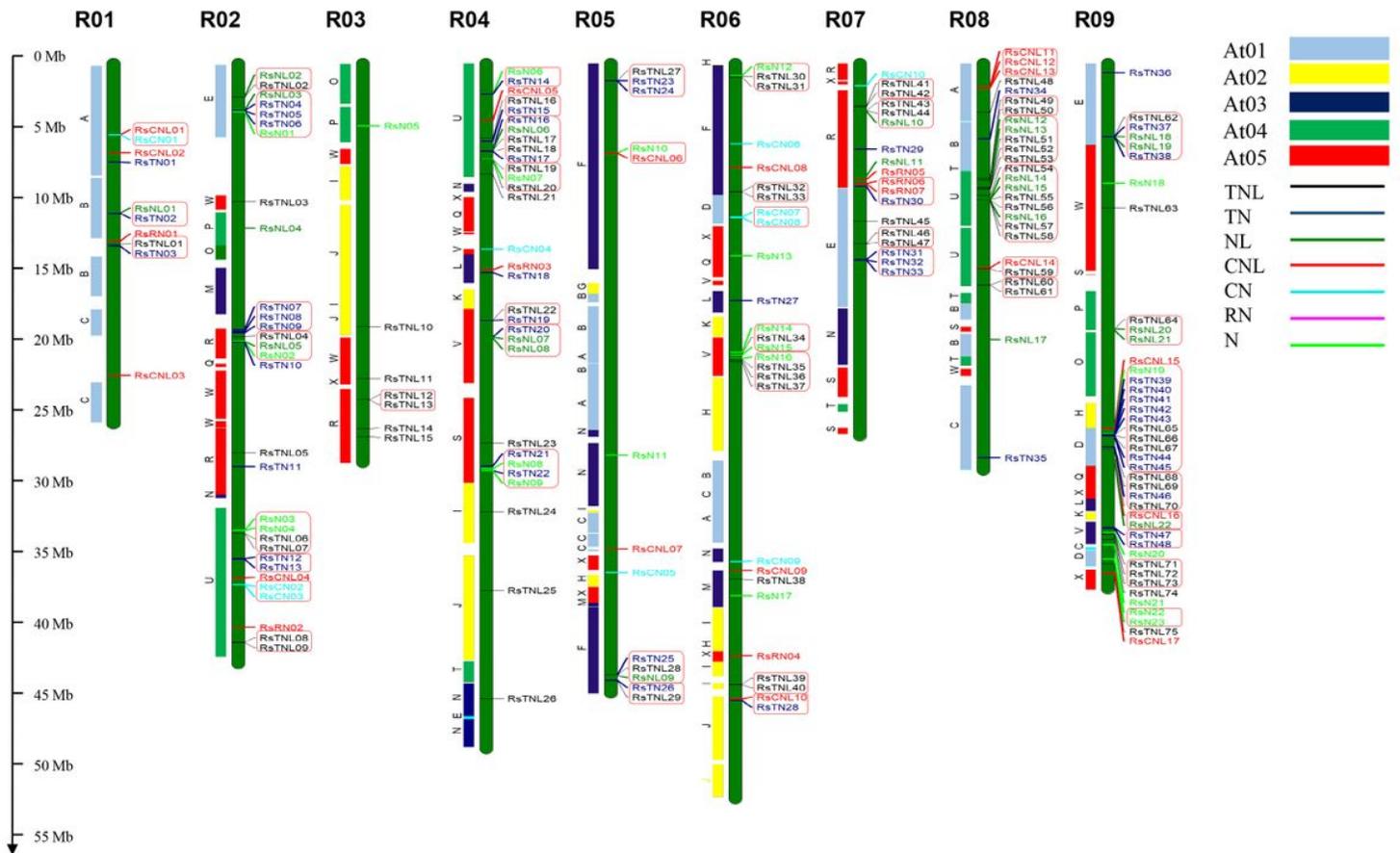
19. Jupe F, Pritchard L, Etherington GJ, MacKenzie K, Cock PJ, Wright F, Sharma SK, Bolser D, Bryan GJ, Jones JD: Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics*. 2012;13(1):75.
20. Alamery S, Tirnaz S, Bayer P, Tollenaere R, Chaloub B, Edwards D, Batley J: Genome-wide identification and comparative analysis of NBS-LRR resistance genes in *Brassica napus*. *Crop and Pasture Science*. 2018;69(1):72-93.
21. Rakow G: Species origin and economic importance of *Brassica*. In: *Brassica*. Springer; 2004: 3-11.
22. Reddy PP: Recent advances in crop protection: Springer Science & Business Media; 2012.
23. Patil S, Sriram S: Biological control of *Fusarium* wilt in crop plants using non-pathogenic isolates of *Fusarium* species. *Indian Phytopathology*.1-9.
24. Pu Z-j, Shimizu M, Zhang Y-j, Nagaoka T, Hayashi T, Hori H, Matsumoto S, Fujimoto R, Okazaki K: Genetic mapping of a *Fusarium* wilt resistance gene in *Brassica oleracea*. *Molecular Breeding*. 2012;30(2):809-818.
25. Okungbowa F, Shittu H: *Fusarium* wilts: An overview. *Environ Res J*. 2012;6(2):122-134.
26. Lv H, Fang Z, Yang L, Zhang Y, Wang Q, Liu Y, Zhuang M, Yang Y, Xie B, Liu B: Mapping and analysis of a novel candidate *Fusarium* wilt resistance gene FOC1 in *Brassica oleracea*. *BMC Genomics*. 2014;15(1):1094.
27. Shimizu M, Pu Z-j, Kawanabe T, Kitashiba H, Matsumoto S, Ebe Y, Sano M, Funaki T, Fukai E, Fujimoto R: Map-based cloning of a candidate gene conferring *Fusarium* yellows resistance in *Brassica oleracea*. *Theoretical and Applied Genetics*. 2015;128(1):119-130.
28. Li J, Chitwood J, Menda N, Mueller L, Hutton SF: Linkage between the I-3 gene for resistance to *Fusarium* wilt race 3 and increased sensitivity to bacterial spot in tomato. *Theoretical and Applied Genetics*. 2018;131(1):145-155.
29. Meyers BC, Vu TH, Tej SS, Ghazal H, Matvienko M, Agrawal V, Ning J, Haudenschild CD: Analysis of the transcriptional complexity of *Arabidopsis thaliana* by massively parallel sigNATURE sequencing. *Nature Biotechnology*. 2004;22(8):1006-1011.
30. Mun J-H, Yu H-J, Park S, Park B-S: Genome-wide identification of NBS-encoding resistance genes in *Brassica rapa*. *Molecular Genetics and Genomics*. 2009;282(6):617.
31. Sagi MS, Deokar AA, Tar'an B: Genetic analysis of NBS-LRR gene family in chickpea and their expression profiles in response to *Ascochyta* blight infection. *Frontiers in Plant Science*. 2017;8:838.
32. Xiang L, Liu J, Wu C, Deng Y, Cai C, Zhang X, Cai Y: Genome-wide comparative analysis of NBS-encoding genes in four *Gossypium* species. *BMC Genomics*. 2017;18(1):292.
33. Mitsui Y, Shimomura M, Komatsu K, Namiki N, Shibata-Hatta M, Imai M, Katayose Y, Mukai Y, Kanamori H, Kurita K: The radish genome and comprehensive gene expression profile of tuberous root formation and development. *Scientific Reports*. 2015;5(1):1-14.
34. Jeong Y-M, Kim N, Ahn BO, Oh M, Chung W-H, Chung H, Jeong S, Lim K-B, Hwang Y-J, Kim G-B: Elucidating the triplicated ancestral genome structure of radish based on chromosome-level

- comparison with the Brassica genomes. *Theoretical and Applied Genetics*. 2016;129(7):1357-1372.
35. Liu Z, Xie J, Wang H, Zhong X, Li H, Yu J, Kang J: Identification and expression profiling analysis of NBS-LRR genes involved in *Fusarium oxysporum* f. sp. *conglutinans* resistance in cabbage. *3 Biotech*. 2019;9(5):202.
36. Richly E, Kurth J, Leister D: Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. *Molecular Biology and Evolution*. 2002;19(1):76-84.
37. Zheng F, Wu H, Zhang R, Li S, He W, Wong F-L, Li G, Zhao S, Lam H-M: Molecular phylogeny and dynamic evolution of disease resistance genes in the legume family. *BMC Genomics*. 2016;17(1):402.
38. McDowell JM, Woffenden BJ: Plant disease resistance genes: recent insights and potential applications. *TRENDS in Biotechnology*. 2003;21(4):178-183.
39. Friedman AR, Baker BJ: The evolution of resistance genes in multi-protein plant resistance systems. *Current Opinion in Genetics & Development*. 2007;17(6):493-499.
40. Cannon SB, Mitra A, Baumgarten A, Young ND, May G: The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biology*. 2004;4(1):10.
41. Yang Y-W, Lai K-N, Tai P-Y, Li W-H: Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineages. *Journal of Molecular Evolution*. 1999;48(5):597-604.
42. Koch MA, Haubold B, Mitchell-Olds T: Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Molecular Biology and Evolution*. 2000;17(10):1483-1498.
43. Panchy N, Lehti-Shiu M, Shiu S-H: Evolution of gene duplication in plants. *Plant Physiology*. 2016;171(4):2294-2316.
44. McHale L, Tan X, Koehl P, Michelmore RW: Plant NBS-LRR proteins: adaptable guards. *Genome Biology*. 2006;7(4):212.
45. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW: Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *The Plant Cell*. 2003;15(4):809-834.
46. Yu X, Kang DH, Choi SR, Ma Y, Lu L, Oh SH, Chhapekar SS, Lim YP: Isolation and characterization of fusarium wilt resistance gene analogs in radish. *3 Biotech*. 2018;8(5):255.
47. Yu X, Lu L, Ma Y, Chhapekar SS, Yi SY, Lim YP, Choi SR: Fine-mapping of a major QTL (Fwr1) for fusarium wilt resistance in radish. *Theoretical and Applied Genetics*. 2019:1-12.
48. Chang A, Lamara M, Wei Y, Hu H, Parkin IA, Gossen BD, Peng G, Yu F: Clubroot resistance gene Rcr6 in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa*. *BMC Plant Biology*. 2019;19(1):224.
49. Finn RD, Clements J, Eddy SR: HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*. 2011;39(suppl\_2):W29-W37.

50. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R: Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947-2948.
51. Lozano R, Hamblin MT, Prochnik S, Jannink J-L: Identification and distribution of the NBS-LRR gene family in the Cassava genome. *BMC Genomics*. 2015;16(1):360.
52. *Bioinformatics Q: CLC Main Workbench*. In.: Release; 2014.
53. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A: The Pfam protein families database in 2019. *Nucleic Acids Research*. 2019;47(D1):D427-D432.
54. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI: CDD: NCBI's conserved domain database. *Nucleic Acids Research*. 2015;43(D1):D222-D226.
55. Bailey TL, Johnson J, Grant CE, Noble WS: The MEME suite. *Nucleic Acids Research*. 2015;43(W1):W39-W49.
56. McDonnell AV, Jiang T, Keating AE, Berger B: Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics*. 2006;22(3):356-358.
57. Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 2004;32(5):1792-1797.
58. Kumar S, Stecher G, Li M, Knyaz C, Tamura K: MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*. 2018;35(6):1547-1549.
59. Whelan S, Goldman N: A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution*. 2001;18(5):691-699.
60. 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 Research*. 2002;30(1):325-327.
61. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H: MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*. 2012;40(7):e49-e49.
62. 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.
63. Gu Z, Cavalcanti A, Chen F-C, Bouman P, Li W-H: Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. *Molecular Biology and Evolution*. 2002;19(3):256-262.
64. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J: KaKs\_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, proteomics & Bioinformatics*. 2010;8(1):77-80.
65. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*. 2009;37(suppl\_2):W202-W208.

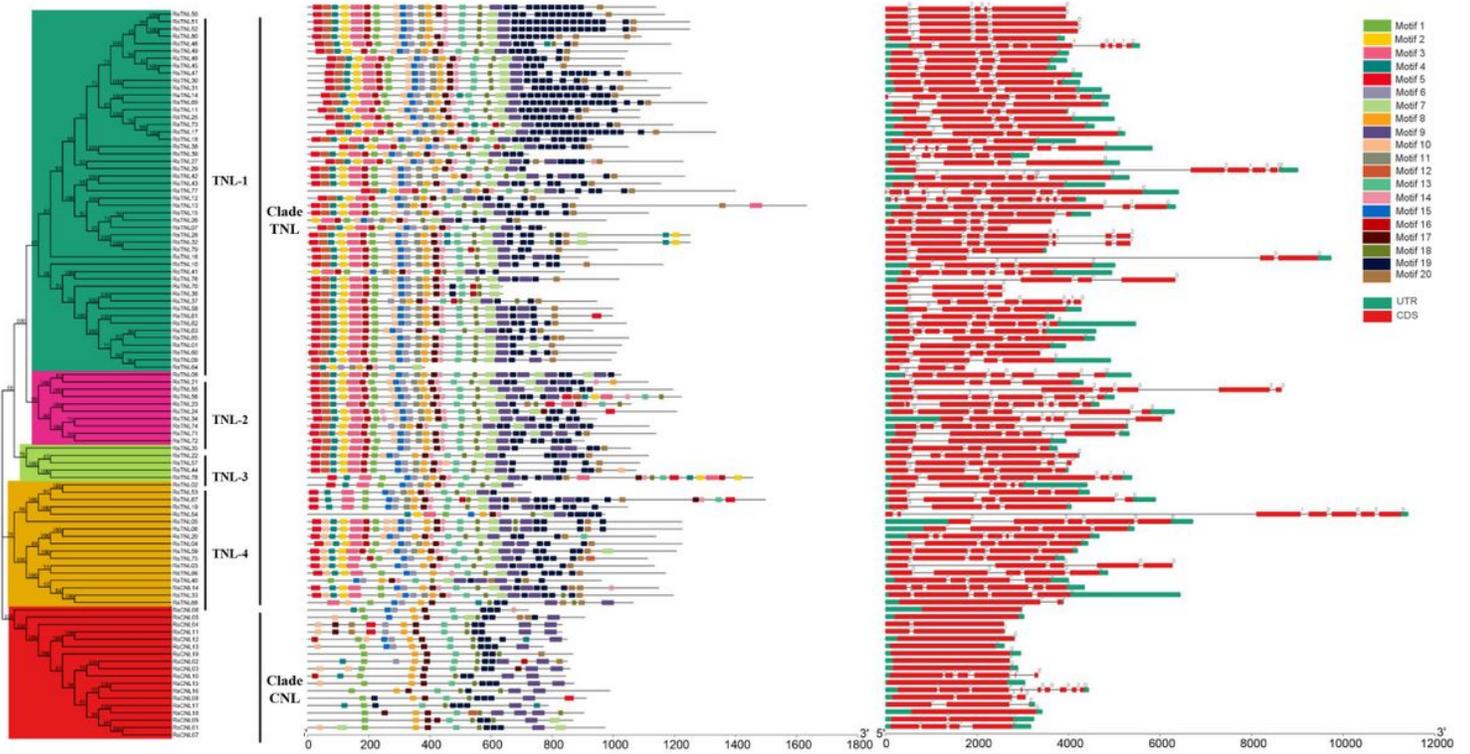


Phylogenetic analysis of the radish NBS-LRR proteins. A maximum likelihood phylogenetic tree was constructed based on 225 NBS-encoding genes. The bootstrap values (1,000 iterations) are provided on each branch. The proteins are designated as follows: Rs+Domain (TNL, TN, NL, CNL, CN, RN, N, and NL). Red and blue correspond to CNL and TNL clades, respectively. The radish NBS-encoding gene ID and the protein sequence information are provided in Additional file 1: Table S1.



**Figure 2**

Chromosomal localization and clustering of NBS-encoding genes in the *R. sativus* genome. The distribution of 202 genes on nine chromosomes (R01–R09) is presented. Different NBS types are marked by different colors. Gene clusters are indicated by red rectangles. The left side of chromosomes represents the syntenic regions in *A. thaliana*. The *A. thaliana* chromosomes are designated as At01–05, which are presented in different colors. The conserved chromosomal blocks in crucifer genomes are indicated by A–X (Schranz et al., 2006).



**Figure 3**

Circos diagram of segmentally duplicated NBS-encoding genes in the radish genome. Gray lines indicate the syntentic blocks in the radish genome. Red lines indicate the duplicated pairs of NBS-encoding genes. Chromosome numbers are provided on each chromosome.



represent untranslated 5' and 3' regions. Red boxes and gray lines indicate exons and introns, respectively. Gene and protein lengths can be estimated with the scale at the bottom.

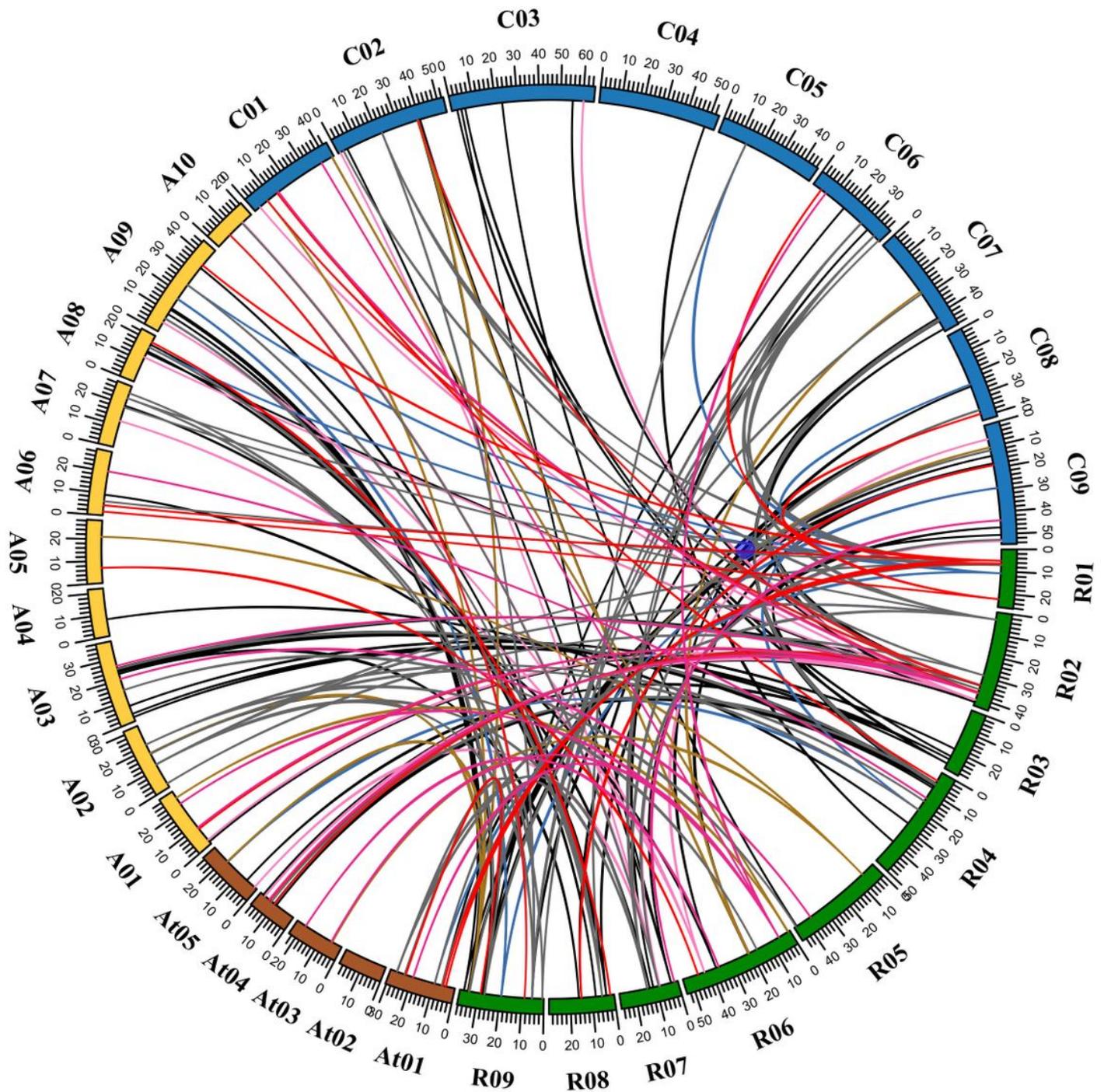
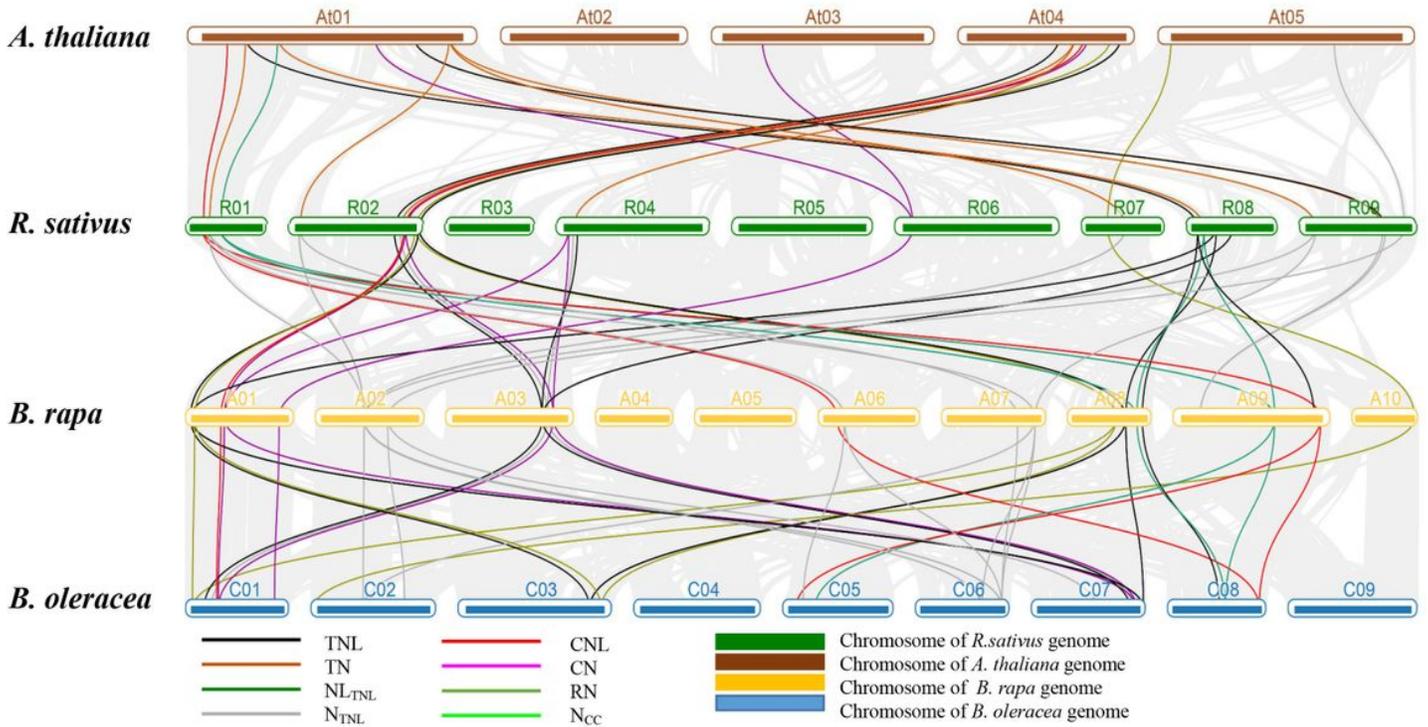


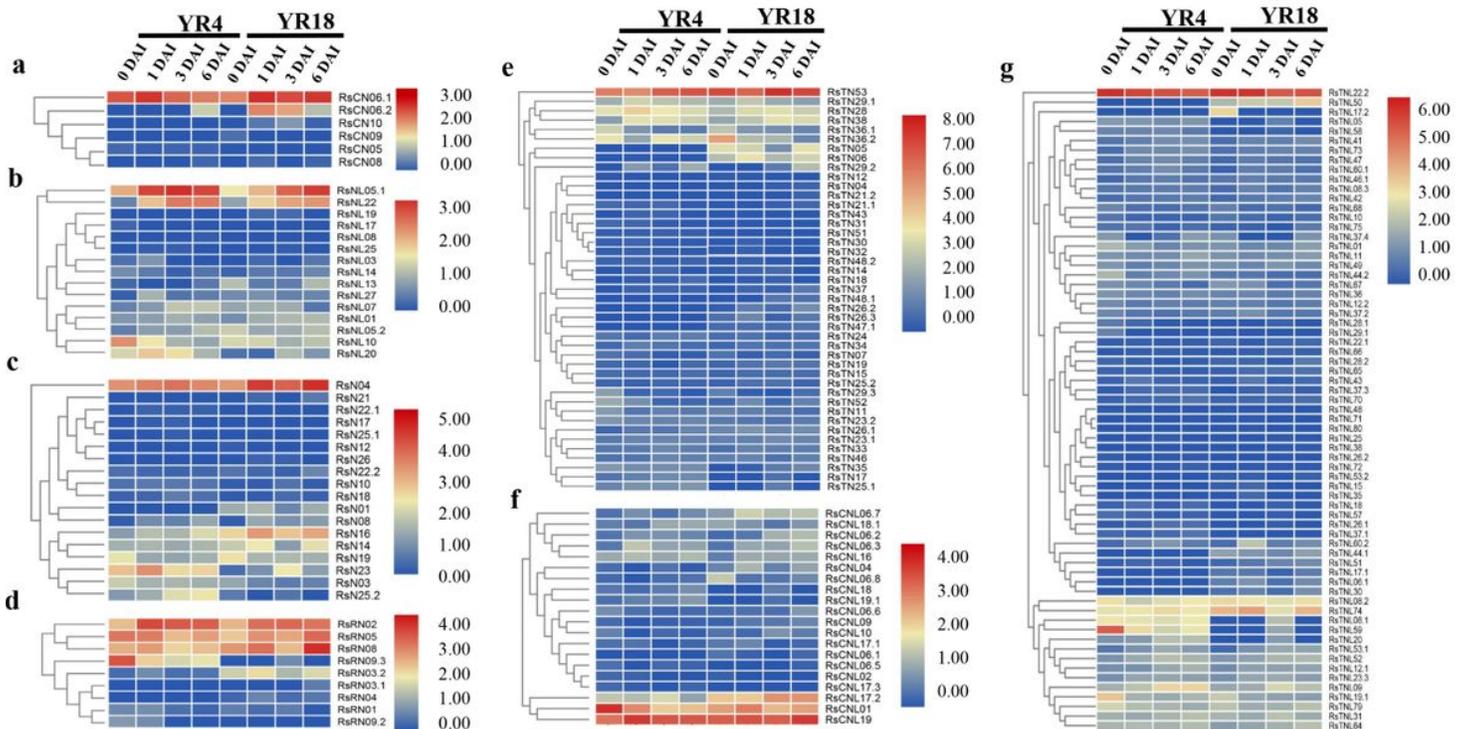
Figure 5

Syntenic relationships between radish NBS-encoding genes and *A. thaliana*, *B. rapa*, and *B. oleracea* genes. Green bars represent *R. sativus* chromosomes (R01–R09). Brown, yellow, and blue bars represent the chromosomes of *A. thaliana* (At01–At05), *B. rapa* (A01–A10), and *B. oleracea* (C01–C09), respectively.



**Figure 6**

Synteny of 22 NBS-encoding genes between radish and *A. thaliana*, *B. rapa*, and *B. oleracea*. Gray lines in the background indicate the collinear blocks within the radish and other genomes. Different colored lines represent the relationships among orthologous gene pairs among the plant species.



**Figure 7**

Expression profiles of NBS-LRR genes in radish. The 'YR4' and 'YR18' plants are resistant and susceptible plant lines, respectively. a Expression of RsCN genes. b Expression of RsNL genes. c Expression of RsN genes. d Expression of RsRN<sub>0</sub> genes. e Expression of RsTN genes. f Expression of RsCNL genes. g Expression of RsTNL genes. DAI: days after inoculation.

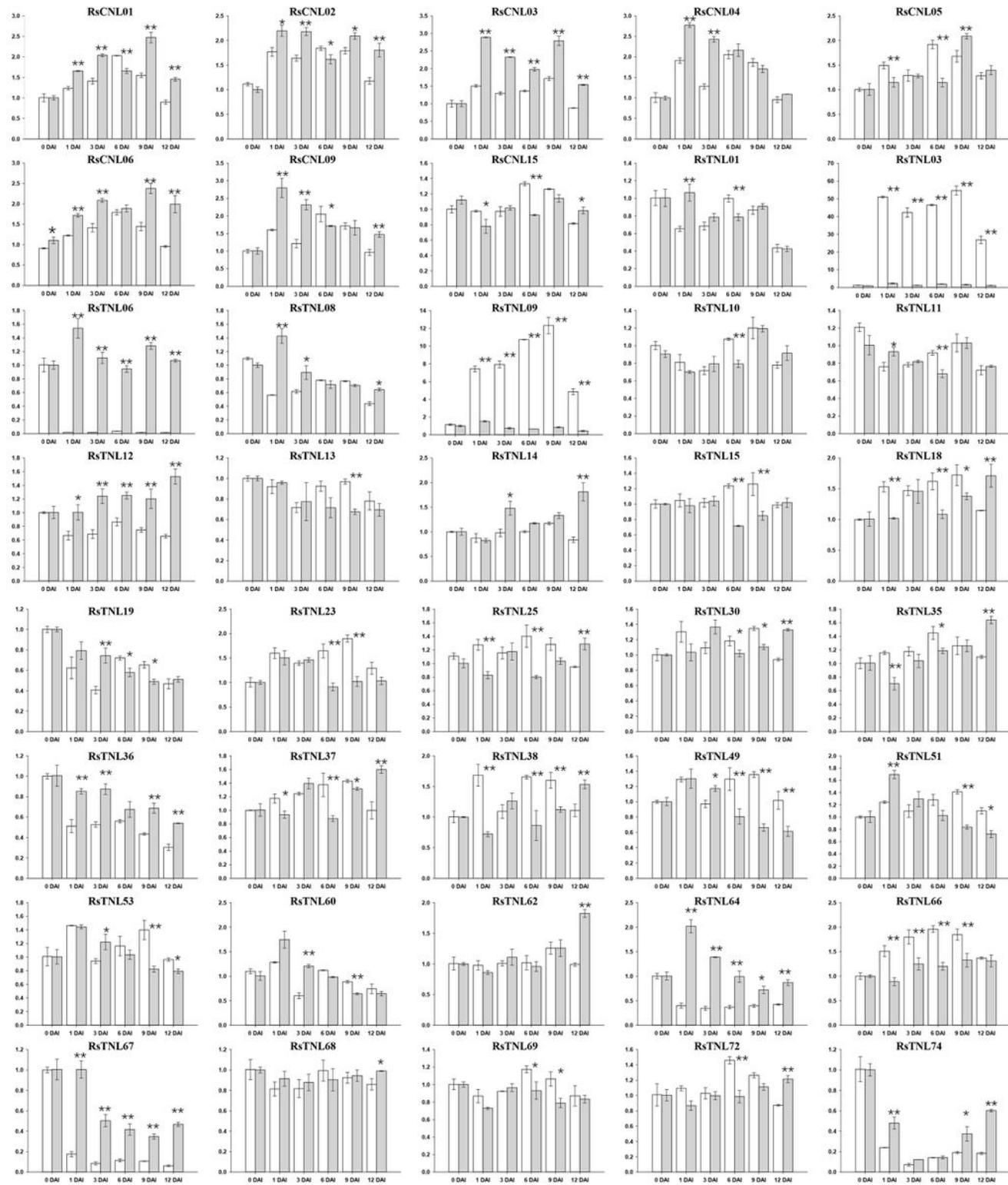


Figure 8

Relative expression levels of 8 RsCNL and 32 RsTNL radish genes. The 'YR4' (resistant) and 'YR18' (susceptible) lines are represented by white and gray bars, respectively. The y-axis represents the relative gene expression levels, whereas the time-points (0, 1, 3, 6, 9 and 12 DAI) are presented on the x-axis. Asterisks indicate significant differences between the 'YR4' and 'YR18' lines based on a t-test.

## Supplementary Files

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

- [Additionalfile3Fig.2.tif](#)
- [Additionalfile2Fig.1.tif](#)
- [Additionalfile1.xlsx](#)