



OPEN ACCESS

EDITED BY

Purushothaman Natarajan,
West Virginia State University,
United States

REVIEWED BY

Joseph Kuhl,
University of Idaho, United States
Krishnanand P. Kulkarni,
Delaware State University, United States

*CORRESPONDENCE

Xiaohui Liang,
✉ 439397013@qq.com

SPECIALTY SECTION

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

RECEIVED 10 January 2023

ACCEPTED 21 February 2023

PUBLISHED 02 March 2023

CITATION

Liang X and Dong J (2023), Comparative-
genomic analysis reveals dynamic NLR
gene loss and gain across
Apiaceae species.
Front. Genet. 14:1141194.
doi: 10.3389/fgene.2023.1141194

COPYRIGHT

© 2023 Liang and Dong. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](#). 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.

Comparative-genomic analysis reveals dynamic NLR gene loss and gain across Apiaceae species

Xiaohui Liang* and Junming Dong

The State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing, Jiangsu, China

Introduction: Nucleotide-binding leucine-rich repeat (NLR) genes play a crucial role in green plants' responding to various pathogens. Genome-scale evolutionary studies of NLR genes are important for discovering and applying functional NLR genes. However, little is known about the evolution of NLR genes in the Apiaceae family including agricultural and medical plants.

Methods: In this study, comparative genomic analysis was performed in four Apiaceae species to trace the dynamic evolutionary patterns of NLR genes during speciation in this family.

Results: The results revealed different number of NLR genes in these four Apiaceae species, namely, *Angelica sinensis* (95), *Coriandrum sativum* (183), *Apium graveolens* (153) and *Daucus carota* (149). Phylogenetic analysis demonstrated that NLR genes in these four species were derived from 183 ancestral NLR lineages and experienced different levels of gene-loss and gain events. The contraction pattern of the ancestral NLR lineages was discovered during the evolution of *D. carota*, whereas a different pattern of contraction after first expansion of NLR genes was observed for *A. sinensis*, *C. sativum* and *A. graveolens*.

Discussion: Taken together, rapid and dynamic gene content variation has shaped evolutionary history of NLR genes in Apiaceae species.

KEYWORDS

Apiaceae, NLR gene, plant disease resistance, evolutionary pattern, gene family

1 Introduction

To cope with the invasion of various pathogenic microorganisms, plants have evolved two layers of immune systems to protect themselves (Wang et al., 2020; Zhou and Zhang, 2020). Relying on the pattern recognition receptors on the surface of plant cells, the first layer immune system helps the plants to recognize conserved pathogen-associated molecular patterns (PAMPs) of bacterial cells, namely, PAMP-triggered immunity (PTI) (Bigeard et al., 2015). Correspondingly, pathogens can block PTI signaling by releasing effector proteins into plant cells (Bigeard et al., 2015). In response to the blocking of bacterial effectors, plants have evolved the second layer immune system that detects the presence of these effectors via intracellular disease resistance genes (R genes) and further induces downstream immunity, namely, effector-triggered immunity (ETI) (Dangl et al., 2013; Wang et al., 2020). Since the first R gene was cloned and characterized in the late 19th century, more than 300 functional R genes have been identified (Johal and Briggs, 1992; Kourelis and van der Hoorn, 2018).

Currently, over 60% of R genes belong to a family that encodes nucleotide-binding site (NBS) and leucine-rich-repeat (LRR) domain receptors, well known as NBS-LRR or NLR genes

(Kourelis and van der Hoorn, 2018). Based on the different N-terminal structural domains, NLR genes were classified into three subclasses, including TNL containing the TIR structural domain, CNL with the CC structural domain, and RNL characterized by the RPW8 structural domain (Whitham et al., 1994; Parker et al., 1997; Botella et al., 1998; Pan et al., 2000; Shao et al., 2014). The majority of CNL and TNL proteins function as detectors of pathogens' effectors (Kourelis and van der Hoorn, 2018), while RNL proteins typically act as "helper" NLRs and get involved in the downstream signaling of CNL and TNL proteins (Kourelis and van der Hoorn, 2018; Wang et al., 2020).

The development of sequencing technology has generated plenty of plant genomes. The genome-scale identification of NLR gene composition not only helps to reveal the evolutionary patterns of NLR gene families in plants, but is also crucial for discovering and applying functional NLR genes, especially for agricultural or industrial crops. Liu et al. (2021) revealed that the convergent contraction of NLR genes was associated with ecological adaptability and NLR subclasses co-evolved with plant immune pathway components. Several studies at the family level have revealed distinctive evolutionary patterns of different plant taxa, such as NLR gene contraction of Poaceae species (Luo et al., 2012), consistent expansion of NLR genes in Fabaceae species (Jia et al., 2015) and first expansion and then contraction of NLR genes in Brassicaceae species (Zhang et al., 2016). Besides, genome-scale NLR gene analysis of multiple wheat lines led to successful cloning of the insect resistance gene *Sml1* (Walkowiak et al., 2020). A highly conserved NLR gene, *Pb3*, was screened based on genome-wide association analysis of 230 rice germplasm genomes and rice blast resistance phenotypes, and further molecular experiments demonstrated that *Pb3* could confer rice blast resistance (Ma et al., 2022).

Apiaceae species are annual or perennial herbs that are widely distributed in temperate regions of the world (Fox et al., 2022). Recently, the genomes of seven species from Apiaceae, including Apioideae and Mackinlayoideae, have been sequenced (https://www.plabipd.de/plant_genomes_pa.ep). Apiaceae species have always played an important role in socio-economic and human life, among which *Angelica sinensis*, *Bupleurum chinense*, *Centella asiatica* and *O. javanica* are medicinal plants (Lu and Li, 2019; Sun et al., 2020; Liu et al., 2022), as well as *Apium graveolens*, *Coriandrum sativum* and *D. carota* are vegetables (Fox et al., 2022). Previous studies have shown that the Ks peak of the paralogous genes in five Apioideae species, including *Daucus carota*, *Oenanthe javanica*, *A. sinensis*, *C. sativum* and *A. graveolens*, corresponded to a recent WGD (whole genome duplication) event, possibly specific in Apioideae instead of Mackinlayoideae (Han et al., 2022). A variety of pathogens are known to be capable of infecting Apiaceae plants and causing economic losses (Fox et al., 2022). However, few studies reported the composition and evolutionary pattern of NLR genes in the Apiaceae family, limiting the discovery of functional NLR genes and their application in production. In this study, the four Apiaceae species from different genera, with fully annotated genomic data, were collected and used to decipher the evolutionary features of NLR genes in this family.

2 Materials and methods

2.1 Identification and classification of NLR genes

Genomic sequences of four Apiaceae species were downloaded from the available database (Table 1). Detailed information related to

the genomes is shown in Table 1. Three Apiaceae species were excluded due to the lack of the gff3 annotation files, namely, *B. chinense*, *O. javanica* and *C. asiatica*. The identification of NLR genes was performed as described previously (Shao et al., 2014). In brief, the NBS (named as NB-ARC) domain was identified using both hidden Markov models search (HMMsearch) and BLAST methods. The HMM profile of the NBS domain (Pfam no. PF00931) was downloaded from the Pfam database and used as a query to search for NLR proteins from four Apiaceae species via hmmer3.3 (Johnson et al., 2010) with the E-value setting to 10^{-4} . To avoid missing possible candidates, the resulting protein sequences were further used to run a BLASTp search against all protein sequences in each genome (E-value = 1.0). To verify whether all hits indeed possessed the NBS domain, they were then subjected to hmmscan analysis using hmmer3.3 (Johnson et al., 2010) against a local Pfam-A database with the E-value setting to 10^{-4} . MEME analysis was conducted to annotate conserved motifs in the NBS domain of all the identified NLR genes (Bailey et al., 2009). Conserved motifs were visualized using WebLogo (Crooks et al., 2004).

2.2 Chromosomal distributions of NLR genes

Chromosomal distribution patterns of NLR genes were identified as described previously (Ameline-Torregrosa et al., 2008). In brief, the genomic positions of obtained NLR genes were extracted from the gff3 annotation files. Then the sliding-window analysis was performed according to the 250 kb window size. An NLR gene was regarded as a singleton locus if no other NLR genes were found upstream or downstream of 250 kb region flanking it. Comparatively, if the distance between two annotated NLR genes was less than 250 kb, they were considered as a clustered locus. Besides, syntenic relationships of NLR genes were analyzed and visualized using Ttools (Chen et al., 2020).

2.3 Phylogenetic analysis of NLR genes

The amino acid sequences of the NBS domains were extracted from all the identified NLR genes. Multiple sequence alignment was performed using ClustalW with default parameters, and then artificially adjusted with MGEA X (Kumar et al., 2018). Those sequences with too short length were removed from the aligned results. The resulting alignments were subjected to the phylogenetic analysis using IQ-TREE with the maximum likelihood method (Minh et al., 2020). The best-fit model of nucleotide substitution was selected by ModelFinder (Kalyaanamoorthy et al., 2017). Branch support values were estimated using SH-aLRT (Anisimova et al., 2011) and UFBoot2 (Minh et al., 2013) with 1,000 bootstrap replicates, and the resulting files were visualized and annotated with iTOL (Letunic and Bork, 2021).

2.4 Gene loss/duplication events of NLR genes

The comparative analysis between phylogenetic tree of NLR genes and species tree was performed using Notung software to

TABLE 1 Metadata of the genomes from four Apiaceae species.

Species	References	Download-link
<i>Angelica sinensis</i>	Han et al. (2022). The Plant Journal	https://data.cyverse.org/dav-anon/iplant/home/licheng_caas/Angelica.sinensis_genome/
<i>Apium graveolens</i>	Cheng et al. (2022). iScience	https://ngdc.cnbc.ac.cn/search/?dbId=gwh&q=GWHBGBL000000000+
<i>Coriandrum sativum</i>	Song et al. (2020). The Plant Journal	http://cgdb.bio2db.com/Download/
<i>Daucus carota</i>	Iorizzo et al. (2016). Nature Genetics	https://www.ncbi.nlm.nih.gov/assembly/GCF_001625215.1

TABLE 2 The number and composition of identified NLR genes in the studied four Apiaceae species.

NLR composition	<i>A. sinensis</i>	<i>C. sativum</i>	<i>A. graveolens</i>	<i>D. carota</i>
NLR	95	183	153	149
CNL	53 (55.79%)	117 (63.93%)	88 (57.52%)	138 (92.62%)
TNL	39 (41.05%)	59 (32.24%)	59 (38.56%)	6 (4.03%)
RNL	3 (3.16%)	7 (3.83%)	6 (3.92%)	5 (3.35%)
Clustered loci	38 (41.76%)	82 (55.03%)	47 (46.08%)	98 (65.77%)
Singleton loci	53 (58.24%)	67 (44.97%)	55 (53.92%)	51 (34.23%)
Unmapped	4	34	51	0

determine gene loss/duplication events of NLR genes (Stolzer et al., 2012). The MCScanX package was used to analyze the types of NLR gene duplication in a given genome according to a pair-wise all-against-all blast of protein sequences (Wang et al., 2012).

3 Results

3.1 Compositions of NLR genes in four Apiaceae species

In this study, the number of NLR genes in the investigated four Apiaceae species showed some variations with ranging from 95 in *A. sinensis* to 183 in *C. sativum* (Table 2). Taken as a reference, the number of NLR genes in the other three species was 1.95 times (*C. sativum*), 1.61 times (*A. graveolens*) and 1.57 times (*D. carota*) as many as that of *A. sinensis*, respectively (Table 2). The subclass classification of NLR genes showed that each of four Apiaceae species possessed all the three subclasses, namely, CNL, TNL and RNL, proposed by previous study (Shao et al., 2016) (Table 2). Among these three NLR subclasses, CNL genes accounted for the highest percentage for each species, reaching 92.62% (*D. carota*), 63.93% (*C. sativum*), 57.52% (*A. graveolens*) and 55.79% (*A. sinensis*), respectively (Table 2). Generally, all Apiaceae species have more CNL and TNL genes than RNL genes, except for *D. carota* which also possessed a small number of TNL genes (Table 2).

Domain analysis showed high structure diversity of NLR proteins among four Apiaceae species (Figure 1). The identified NLR proteins could be divided into 12 groups based on their domain composition and arrangement (Figure 1A). Among them, only most TNL genes could encode complete TNL proteins containing the N-terminal TIR domain, the central NBS domain and the

C-terminal LRR domain (Figure 1A). However, most CNL genes could encode incomplete CNL proteins lacking the N-terminal CC domain, or the C-terminal LRR domain, or both (Figure 1A). A similar scenario was also observed for most RNL genes. The trends in four Apiaceae species were overall consistent. For MEME analysis, four key motifs, namely, P-loop, Kinase-2, RNBS-B and GLPL, were detected in the amino acid sequences of NBS domains from four Apiaceae species (Figure 1B). These four key motifs were found to be conserved in NLR proteins of four Apiaceae species (Figure 1B), as reported in other angiosperms (Shao et al., 2016).

3.2 Chromosomal distribution patterns of NLR genes in four Apiaceae species

To explore the clustering organization, we further analyzed the chromosomal distribution patterns of NLR genes in the four Apiaceae species. The results demonstrated that the majority of NLR genes could be mapped to the specific chromosomes for each species (Figure 2A), except for four (4.21%) in *A. sinensis*, 34 (18.58%) in *C. sativum*, 51 (33.33%) in *A. graveolens* and zero (0%) in *D. carota* due to that these NLR genes were annotated only at the scaffold level (Table 2). Besides, an uneven gene distribution was found among different chromosomes in each species (Figure 2A). When the chromosome with a minimum number of NLR genes was used as a reference, the times of NLR gene on other chromosomes ranged from two to 18 in *A. sinensis*, from two to 15 in *C. sativum*, from two to 26 in *A. graveolens*, and from one to six in *D. carota* (Figure 2A).

As to the clustering situations, the majority of NLR genes belonged to clustered loci rather than singleton loci in two species, namely, *C. sativum* with 55.03% and *D. carota* with 65.77% (Table 2). However, the number of NLR genes treated as singleton loci was

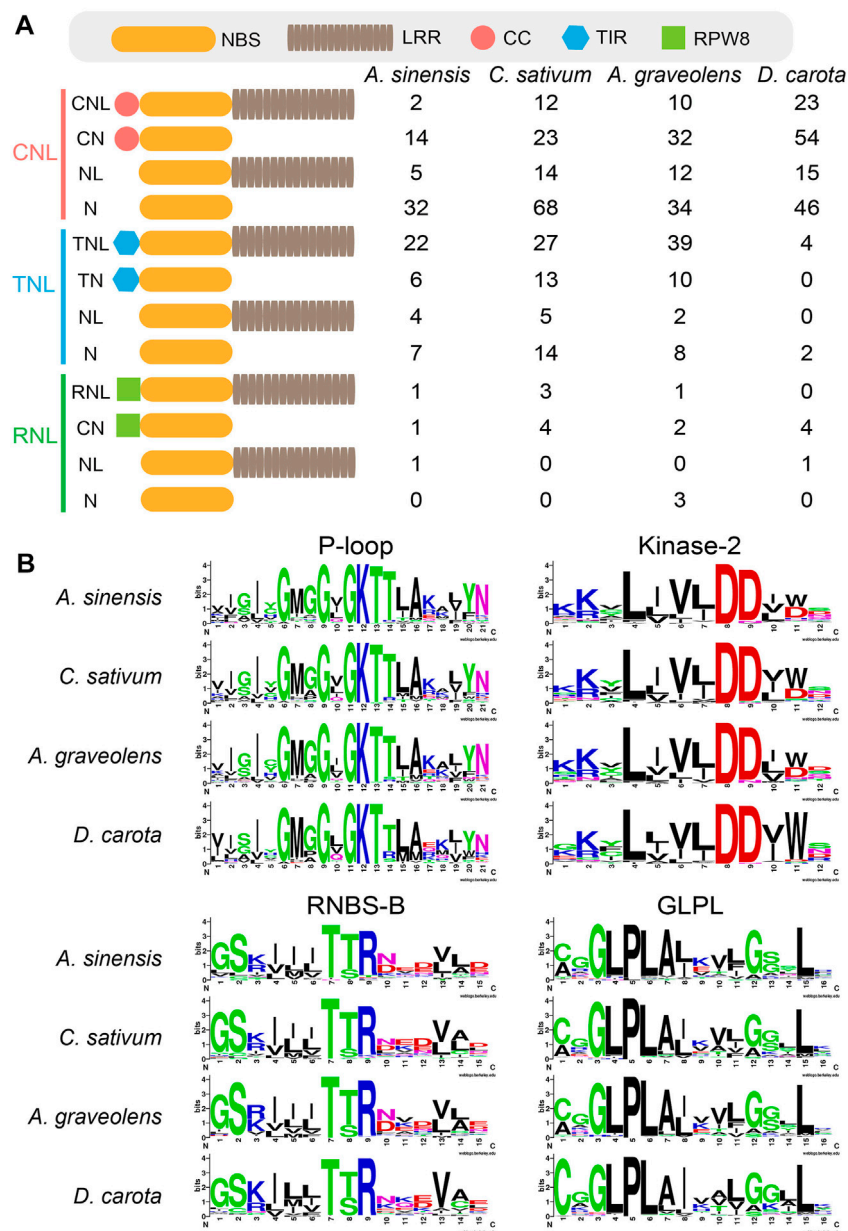


FIGURE 1 Domain structures and key motifs of NLR genes in four Apiaceae species. (A) Domain compositions and arrangements of NLR proteins encoded by the identified NLR genes from four Apiaceae species. C, CC domain; T, TIR domain; R, RPW8 domain; N, NBS domain; L, LRR domain. (B) Four key motifs in the conserved NBS domains from four Apiaceae species.

slightly more than that of NLR genes classified as clustered loci in the remaining two species (Table 2). Besides, the number of clustered loci showed different trends among species, with the highest number in *D. carota* (98), followed by *C. sativum* (82), *A. graveolens* (47) and *A. sinensis* (38) (Table 2). When considering the chromosomal scale, clustered loci were found in almost all chromosomes with the high number of NLR genes (Figure 2A) and the percentage of clustered loci showed some differences among different chromosome for each of these four Apiaceae species, ranging from 22.22% to 66.67% in *A. sinensis*, 33.33%–77.42% in *C. sativum*, 37.50%–69.23% in *A. graveolens* and 25.00%–88.89% in *D. carota* (Figure 2B).

The emergence of clustered loci of NLR genes may be mainly attributed to the occurrence of duplication events (Leister, 2004). Therefore, the duplication types of NLR genes were further investigated in these four Apiaceae species. The results showed that the expansion of NLR genes was dominated by different duplication types among species (Figure 3A). Most of the NLR genes in *A. sinensis* (63.74%, 58/91), *C. sativum* (59.73%, 89/149), *A. graveolens* (50%, 51/102) and *D. carota* (38.93%, 58/149) were resulted from dispersed duplication, whereas the moderate percentage of NLR genes in these four Apiaceae species were characterized as tandem/proximal duplications, except that the

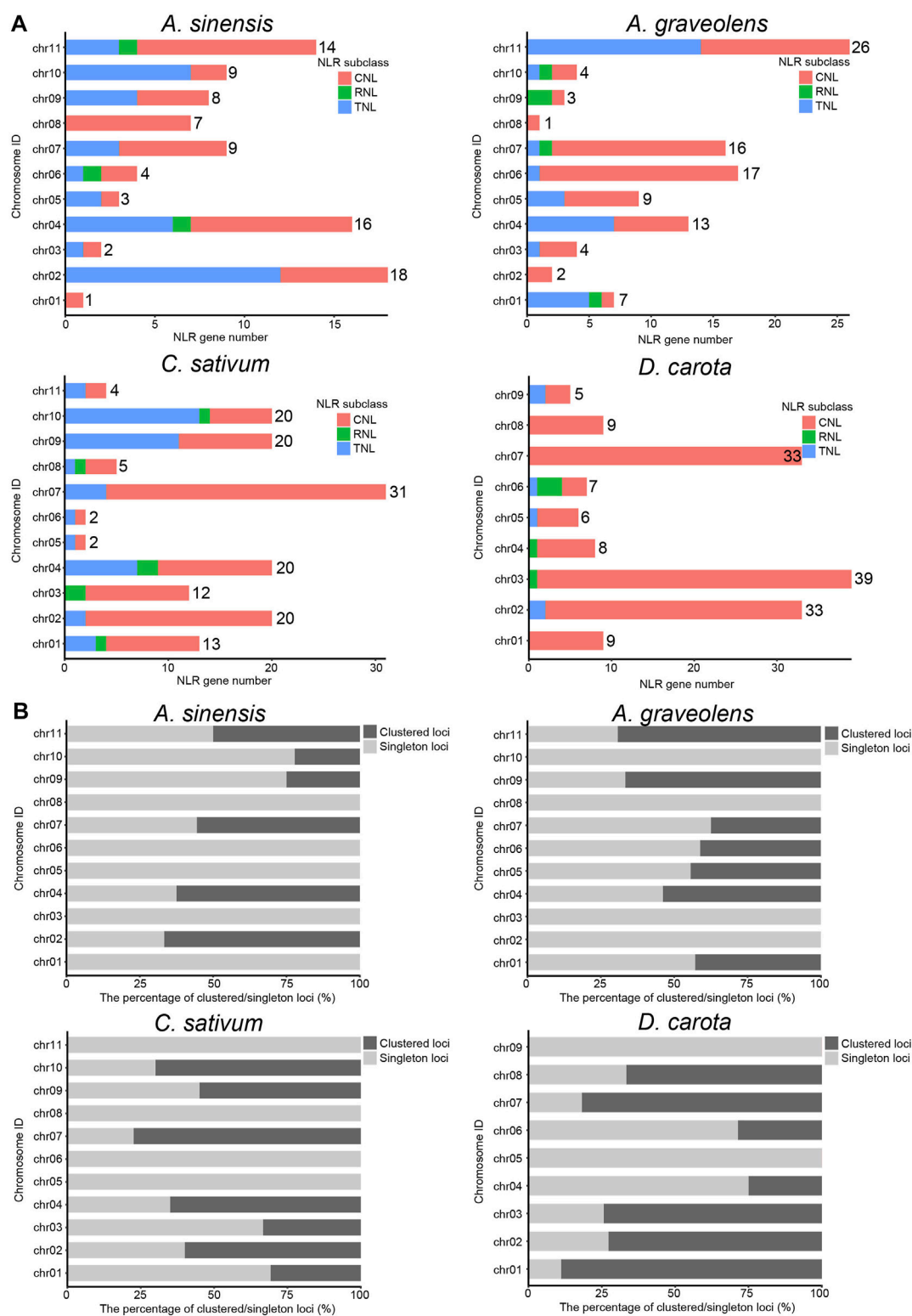


FIGURE 2

The composition and chromosomal distribution of NLR genes in four Apiaceae species. (A) The number variation of NLR genes among different genomes from four Apiaceae species. Three NLR subclasses are indicated with different colors, with red for CNL, green for RNL and blue for TNL. The number of NLR genes on each chromosome is shown. (B) The percentage of clustered/singleton loci among different genomes from four Apiaceae species.

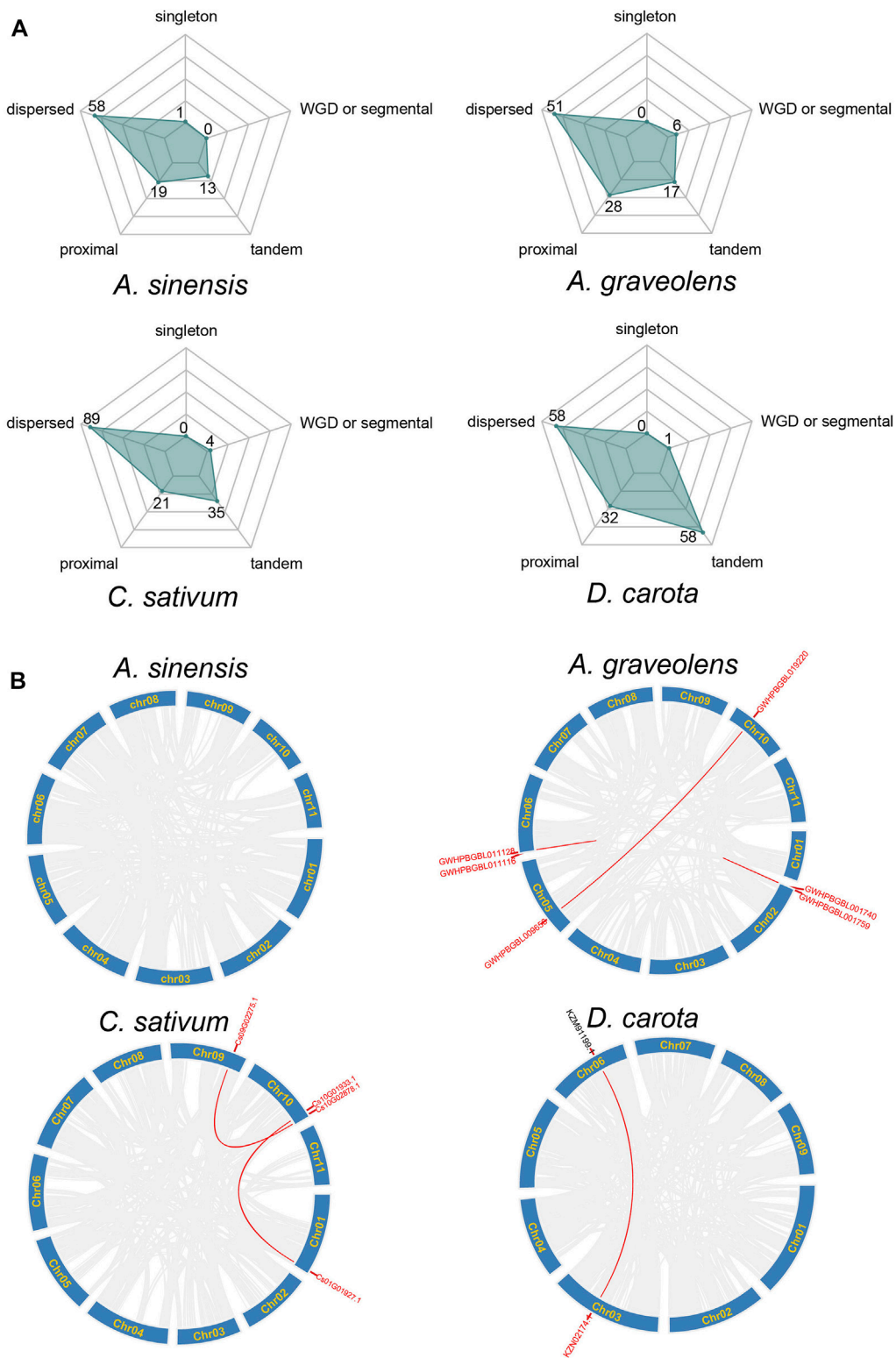


FIGURE 3 Duplication types of NLR genes across Apiaceae species. **(A)** Type of NLR gene duplication in *A. sinensis*, *A. graveolens*, *C. sativum* and *Daucus carota*, respectively. The number of NLR genes derived from each type was shown. **(B)** Syntenic relationships of NLR genes in *A. sinensis*, *A. graveolens*, *C. sativum* and *Daucus carota*, respectively. Gene ID in red represents NLR genes and gene ID in dark belongs to non-NLR gene.

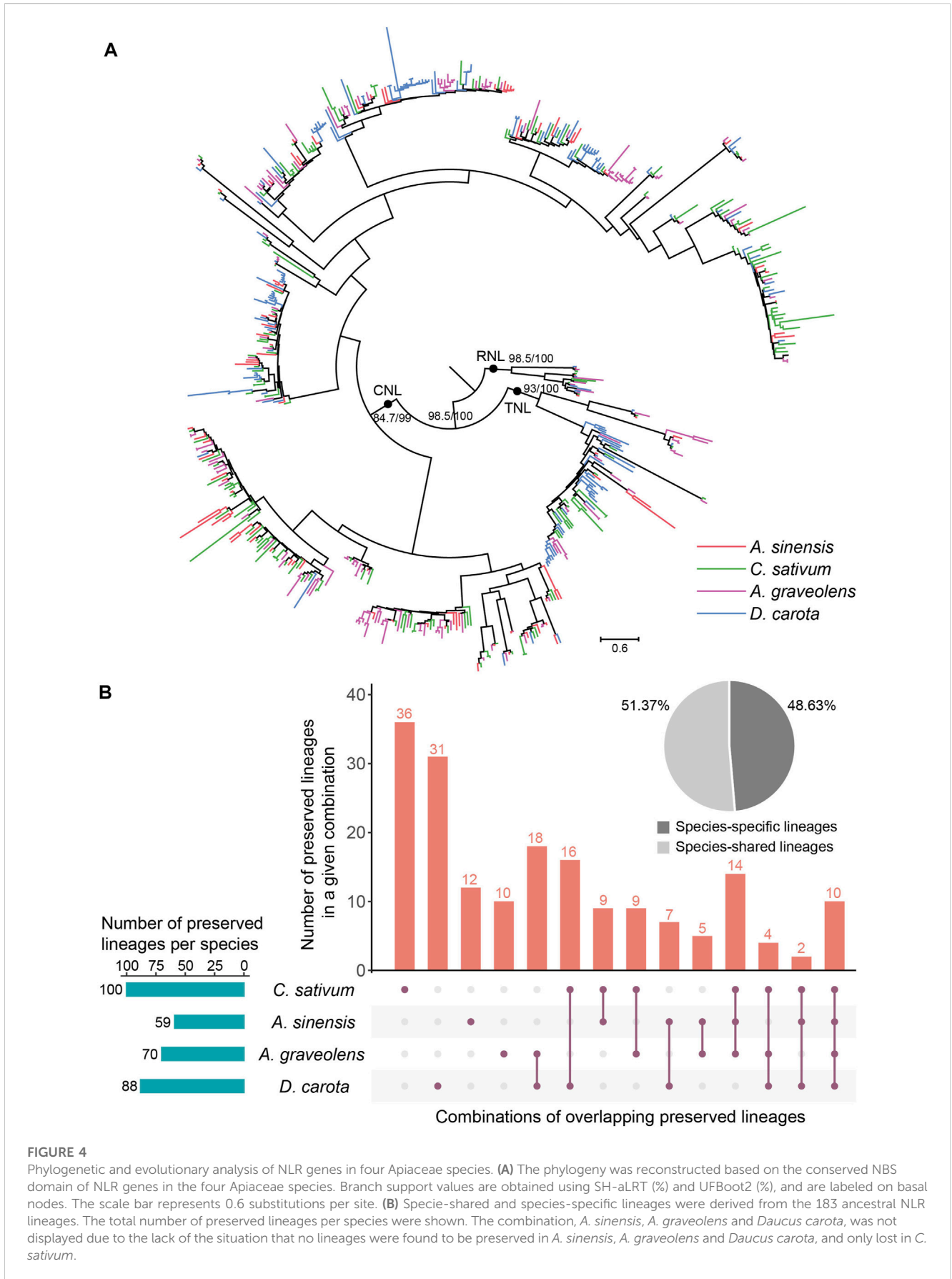


FIGURE 4

Phylogenetic and evolutionary analysis of NLR genes in four Apiaceae species. (A) The phylogeny was reconstructed based on the conserved NBS domain of NLR genes in the four Apiaceae species. Branch support values are obtained using SH-aLRT (%) and UFBoot2 (%), and are labeled on basal nodes. The scale bar represents 0.6 substitutions per site. (B) Specie-shared and species-specific lineages were derived from the 183 ancestral NLR lineages. The total number of preserved lineages per species were shown. The combination, *A. sinensis*, *A. graveolens* and *Daucus carota*, was not displayed due to the lack of the situation that no lineages were found to be preserved in *A. sinensis*, *A. graveolens* and *Daucus carota*, and only lost in *C. sativum*.

NLR gene number *via* tandem duplication events was consistent with that derived by dispersed duplication in *D. carota* (Figure 3A). Only a small percentage of NLR genes in *A. graveolens* (5.88%, 6/102), *C. sativum* (2.68%, 4/149) and *D. carota* (0.67%, 1/149) were generated by whole genome duplications (WGD)/segmental duplication, whereas no WGD/segmental duplicated NLR genes were found in *A. sinensis* (Figure 3A). Correspondingly, the syntenic relationships among NLR genes, implying possible WGD events, were discovered in *A. graveolens*, *C. sativum* and *D. carota*, and no syntenic NLR gene was identified in *A. sinensis* (Figure 3B). Notably, the WGD or segmental duplication event was detected between a RNL gene (KZN02174.1) and a gene (KZM91199.1) only encoding RPW8 domain in *D. carota* (Figure 3B).

3.3 Phylogenetic analysis of NLR genes

To trace the evolutionary history of NLR genes in Apiaceae, their phylogenetic relationships were constructed based on the amino acid sequences of the conserved NBS domain (Figure 4A). The results showed that NLR genes of the studied four Apiaceae species formed three distinct clades with high support values (>80), corresponding to CNL, TNL and RNL subclasses, respectively (Figure 4A). Compared to the CNL and TNL clades, the branch lengths of the RNL clade were relatively shorter (Figure 4A), suggesting that RNL genes had a slower evolutionary rate. In addition, within the CNL and TNL clades, clustering situations of CNL genes and/or TNL genes from a single species were frequently observed (Figure 4A), possibly attributed to species-specific expansions of CNL and/or TNL genes.

To understand the evolutionary patterns of NLR genes during the speciation of these four Apiaceae species, comparative analysis of species tree and gene tree based on NBS domain was performed using Notung software (Stolzer et al., 2012). The results revealed that the common ancestor of these four Apiaceae species possessed at least 183 ancestral NLR lineages, including 143 CNL, 32 TNL and eight RNL (Supplementary Figure S1). These ancestral NLR lineages were further termed as “Apiaceae NLR lineages”. At the lineage level, species-specific and ancestral loss events were found in these four Apiaceae species (Figure 4B). By assigning the NLR genes in each species to the 183 ancestral Apiaceae NLR lineages, the results showed that 59, 100, 70 and 88 Apiaceae NLR lineages were inherited by *A. sinensis*, *C. sativum*, *A. graveolens* and *D. carota*, respectively (Figure 4B). Among these ancestral NLR lineages, only 10 of them were preserved in all the four genomes (Supplementary Table S1). In contrast, 89 lineages were species-specific (*A. sinensis*, 12; *C. sativum*, 36; *A. graveolens*, 10; *D. carota*, 31), reaching 48.63% of the Apiaceae NLR lineages (Figure 4B). The remaining 94 lineages were differentially preserved in two to three species, accounting for 51.37% of the Apiaceae NLR lineages (Figure 4B). Notably, no lineages were found to be preserved in *A. sinensis*, *A. graveolens* and *D. carota*, and only lost in *C. sativum*.

3.4 Evolutionary trajectory of NLR genes in Apiaceae species

Based on the loss and gain events of NLR genes in each divergence node of Apiaceae species, the evolutionary trajectories

of NLR genes in these four Apiaceae species were traced. The 183 ancestral Apiaceae NLR lineages were identified in the common ancestor of these four Apiaceae species (Figure 5A). After separation from the common ancestor of *A. graveolens*, *C. sativum* and *A. sinensis* (Ag-Cs-As node), 95 gene-loss and 56 gain events were detected during the evolution of *D. carota* (Figure 5A), suggesting a pattern of NLR contraction during this period (Figure 5B). However, a different pattern of “first expansion and then contraction” of NLR genes was found for the remaining three species, namely, *A. sinensis* (As), *C. sativum* (Cs) and *A. graveolens* (Ag) (Figure 5B). The expansion events were only detected during the period from the common ancestor of these four Apiaceae species to the Ag-Cs-As node, with 81 gene-gain and only 31 gene-loss events (Figure 5A). After the above-mentioned expansion events, the successive contraction pattern was discovered during the period that the Ag-Cs-As node further diverged into *A. graveolens* (Ag), *C. sativum* (Cs) and *A. sinensis* (As) (Figure 5B). Although different gene-loss and gene-gain events were both found, the number of gene-loss events was more than that of gene-gain events during each evolutionary period since the Ag-Cs-As node (Figure 5A). After separation from the common ancestor of *C. sativum* and *A. sinensis* (Cs-As node), 44 gene-loss and only nine gain events were detected during the evolution of *A. graveolens* (Figure 5A), suggesting a contraction pattern of NLR genes (Figure 5B). The Cs-As node then diverged to generate *A. sinensis* and *C. sativum*. In *A. sinensis*, 115 gene-loss and only 11 gain events were discovered, leading to the existence of 94 NLR genes (Figure 5A). Using this strategy, a similar pattern of “consistent contraction” was also observed for NLR genes in *C. sativum*, with 59 gene-loss and 39 gene-gain events (Figure 5A).

4 Discussion

Shao et al. (2019) suggest that NLR genes originated prior to the last common ancestor of green plants, due to the presence of NLR genes in all the lineages of green plants, as well as the absence in Rhodophyta and Glaucophyta. The NLR genes are well known as a large gene family in angiosperms, with an average of several hundred members per species (Liu et al., 2021). Systematic genome-wide studies of NLR genes have greatly contributed to discovering and applying functional NLR genes in important crops and economic plants in recent years (Walkowiak et al., 2020; Ma et al., 2022). However, the lack of informative evolution on NLR genes has heavily hampered identification and application of functional NLR genes in Apiaceae species. Taking advantage of available genomic data from the four Apiaceae species, the evolutionary profile of NLR genes was analyzed from multiple perspectives in this study, which can provide scientific guidance for breeding of the Apiaceae species.

The intensive genome-scale studies have shown that the number of NLR genes varied dramatically across different species in recent years (Xue et al., 2020; Liu et al., 2021; Santos et al., 2022). For example, both *Apostasia shenzhenica* and *Phalaenopsis equestris* have less than 100 NLR genes (Xue et al., 2020), whereas barley possesses 468 NLR genes (Li et al., 2021) and *Coffea eugenioides* has more than 1000 NLR genes (Santos et al., 2022). In this study, a total of 95, 183, 153 and 149 NLR genes were identified in *A. sinensis*, *C.*

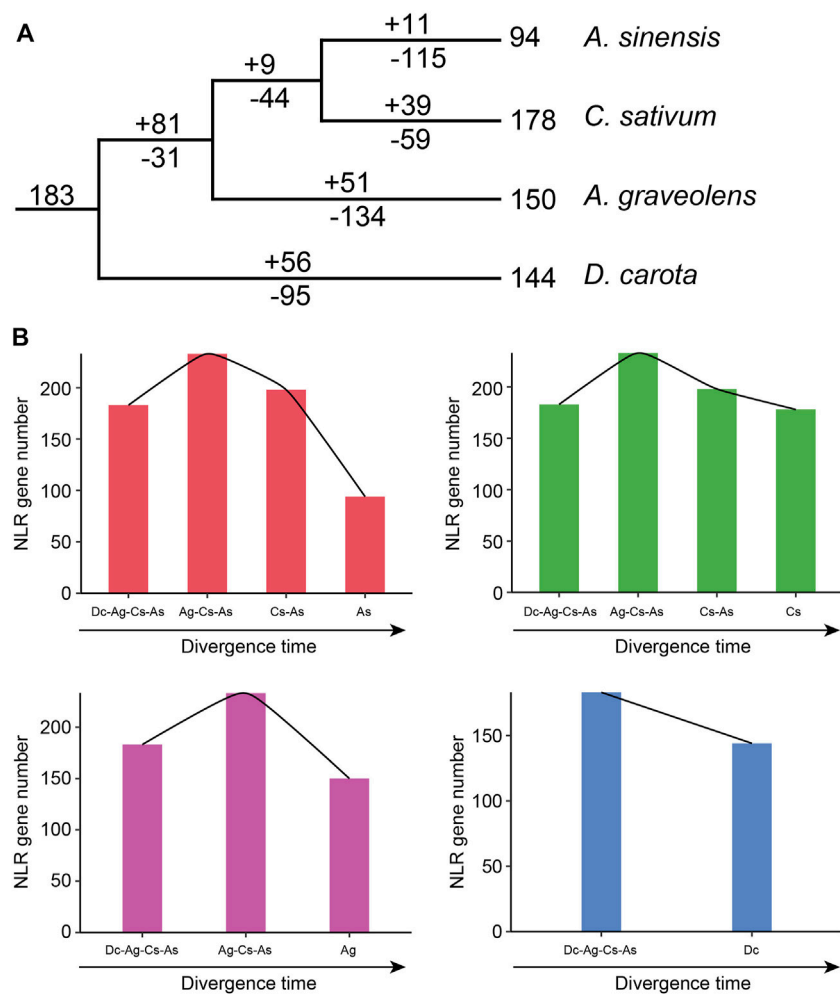


FIGURE 5

Evolutionary trajectory of NLR genes during Apiaceae species speciation. (A) Number variations of NLR genes at different stages of four Apiaceae species. Differential gene-gain and loss events are indicated by numbers with + or - on each branch. (B) Dynamic evolutionary patterns of NLR genes at different stages of four Apiaceae species. As, *A. sinensis*; Cs, *C. sativum*; Ag, *A. graveolens*; Dc, *Daucus carota*.

sativum, *A. graveolens* and *D. carota*, respectively (Table 2). Among all investigated angiosperms up to now, the Apiaceae species were slightly poor in NBS genes, compared to the average of 300 NLR genes in angiosperms (Liu et al., 2021). However, some differences were detected for the number of NLR genes among these four Apiaceae species. The genomes of *C. sativum*, *A. graveolens* and *D. carota* had more than 1.5 times the number of NLR genes as *A. sinensis*, suggesting that species-specific gene-loss and gain events occurred after their separation from the common ancestor. Even if a similar number of NLR genes was found between *A. graveolens* and *D. carota*, significantly different compositions of NLR genes were identified among them (Table 2). This might be attributed to the differential inheritance of ancestral NLR lineages from the common ancestor in these two species (Figure 4B, Figure 5A).

Reconstructing the ancestral states of NLR genes showed that the common ancestor of these four Apiaceae species likely possessed at least 183 ancestral NLR lineages, including 143 CNL, 32 TNL and eight RNL (Supplementary Figure S1). The number of ancestral NLR

lineages in Apiaceae is larger than that in Fabaceae (119, Shao et al., 2014) and Solanaceae (176, Qian et al., 2017), and slightly fewer than that in Brassicaceae (228, Zhang et al., 2016). Besides, the number of ancestral NLR lineages in Apiaceae is much larger than the ancestral NLR lineage number in several monocot groups, such as 30 ancestral lineages in orchids (Xue et al., 2020) and 101 ancestral NLR lineages in Arecaceae (Li et al., 2021). The above-mentioned dicot families had larger ancestral NLR lineage numbers than that in several monocot families, attributed in part to their preserved an additional NLR subclass, TNL. Notably, 10 ancestral NLR lineages were found readily preserved in these four Apiaceae species (Figure 4B), representing four different genera and inhabiting different environments. Those genes in each species, belonging to the 10 ancestral NLR lineages, were available in Supplementary Table S1 and could be potential candidates for the discovery and experimental validation of functional NLR genes. Besides, the absence of orthologous genes in closely related species within the same family may

indicate loss of plant defense against specific pathogens. On this regard, specific probes/markers can be designed to validate the association of presence/absence of NLR genes with plant resistant to various pathogens of Apiaceae species in future studies.

The number of NLR genes in plants has undergone a rapid and dynamic evolution through frequent gene-gain and gene-loss events in response to the fluctuations of external pathogenic environment (Shao et al., 2019; Liu et al., 2021). For example, the Cucurbitaceae and Poaceae species experienced a similar contraction pattern of NLR genes (Li et al., 2010; Luo et al., 2012; Wan et al., 2013), whereas those species in Fabaceae and Rosaceae exhibited a different pattern of NLR gene “consistent expansion” (Shao et al., 2014; Jia et al., 2015). In this study, two different evolutionary modes of NLR genes were revealed in these four Apiaceae species after their diverging from the common ancestor. A distinct “first expansion and then contraction” pattern of NLR genes was observed in the three Apiaceae species (Figure 5), consistent with previous reports in the five Brassicaceae species (Zhang et al., 2016), and some degree of contraction pattern was discovered for *D. carota* (Figure 5).

5 Conclusion

This study systematically analyzed the profiles of NLR genes in four Apiaceae species and revealed their compositions and evolutionary modes, which might provide a resource for discovering and applying functional NLR genes. Phylogenetic analysis showed that NLR genes in these four Apiaceae species were derived from 183 ancestral lineages, and two distinct patterns, “first expansion and then contraction” and “consistent contraction”, were observed for NLR genes in these four species. Species-specific lineage loss and gene duplication have shaped the current NLR profiles of different Apiaceae species. The obtained results can serve as a valuable resource for genetic breeding of Apiaceae species.

References

- Ameline-Torregrosa, C., Wang, B. B., O’Bleness, M. S., Deshpande, S., Zhu, H. Y., Roe, B., et al. (2008). Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiol.* 146, 5–21. doi:10.1104/pp.107.104588
- Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C., and Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* 60, 685–699. doi:10.1093/sysbio/syr041
- 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, W202–W208. doi:10.1093/nar/gkp335
- Bigeard, J., Colcombet, J., and Hirt, H. (2015). Signaling mechanisms in pattern-triggered immunity (PTI). *Mol. Plant* 8, 521–539. doi:10.1016/j.molp.2014.12.022
- Botella, M. A., Parker, J. E., Frost, L. N., Bittner-Eddy, P. D., Beynon, J. L., Daniels, M. J., et al. (1998). Three genes of the arabidopsis RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. *Plant Cell.* 10, 1847–1860. doi:10.1105/tpc.10.11.1847
- Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi:10.1016/j.molp.2020.06.009
- Cheng, Q., Sun, L., Qiao, H., Li, Z. X., Li, M. X., Cui, X. Y., et al. (2022). Loci underlying leaf agronomic traits identified by re-sequencing celery accessions based on an assembled genome. *iScience* 25 (7), 104565. doi:10.1016/j.isci.2022.104565
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Res.* 14, 1188–1190. doi:10.1101/gr.849004
- Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi:10.1126/science.1236011
- Fox, A., Gibbs, A. J., Fowkes, A. R., Pufal, H., McCreig, S., Jones, R. A. C., et al. (2022). Enhanced apiaceous potyvirus phylogeny, novel viruses, and new country and host records from sequencing Apiaceae samples. *Plants-Basel* 11, 1951. doi:10.3390/plants11151951
- Han, X. X., Li, C., Sun, S. C., Ji, J. J., Nie, B., Maker, G., et al. (2022). The chromosome-level genome of female ginseng (*Angelica sinensis*) provides insights into molecular mechanisms and evolution of coumarin biosynthesis. *Plant J.* 112(5):1224–1237. doi:10.1111/tj.16007
- Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J. Y., et al. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet.* 48 (6), 657–666. doi:10.1038/ng.3565
- Jia, Y. X., Yuan, Y., Zhang, Y. C., Yang, S. H., and Zhang, X. H. (2015). Extreme expansion of NBS-encoding genes in Rosaceae. *Bmc Genet.* 16, 48. doi:10.1186/s12863-015-0208-x
- Johal, G. S., and Briggs, S. P. (1992). Reductase-activity encoded by the *Hm1* disease resistance gene in maize. *Science* 258, 985–987. doi:10.1126/science.1359642

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

XL conceived and designed the study. JD obtained and analyzed the data. XL wrote and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

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.2023.1141194/full#supplementary-material>

- Johnson, L. S., Eddy, S. R., and Portugaly, E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinforma.* 11, 431. doi:10.1186/1471-2105-11-431
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi:10.1038/nmeth.4285
- Kourelis, J., and van der Hoorn, R. A. L. (2018). Defended to the nines: 25 Years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell.* 30, 285–299. doi:10.1105/tpc.17.00579
- 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, 1547–1549. doi:10.1093/molbev/msy096
- Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet.* 20, 116–122. doi:10.1016/j.tig.2004.01.007
- Leticia, I., and Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi:10.1093/nar/gkab301
- Li, J., Ding, J., Zhang, W., Zhang, Y. L., Tang, P., Chen, J. Q., et al. (2010). Unique evolutionary pattern of numbers of graminaceous NBS-LRR genes. *Mol. Genet. Genomics* 283, 427–438. doi:10.1007/s00438-010-0527-6
- 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.* 12, 694682. doi:10.3389/fgene.2021.694682
- Li, X.-T., Zhou, G.-C., Feng, X.-Y., Zeng, Z., Liu, Y., and Shao, Z.-Q. (2021). Frequent gene duplication/loss shapes distinct evolutionary patterns of NLR genes in Arecaceae species. *Horticulturae* 7, 539. doi:10.3390/horticulturae7120539
- Liu, L. S., Li, H. P., Tan, G. S., and Ma, Z. J. (2022). Traditional Chinese herbal medicine in treating amenorrhea caused by antipsychotic drugs: Meta-analysis and systematic review. *J. Ethnopharmacol.* 289, 115044. doi:10.1016/j.jep.2022.115044
- 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, 2015–2031. doi:10.1016/j.molp.2021.08.001
- Lu, C. L., and Li, X. F. (2019). A review of *Oenanthe javanica* (blume) DC. As traditional medicinal plant and its therapeutic potential. *Evidence-Based Complementary Altern. Med.* 2019, 1–17. doi:10.1155/2019/6495819
- Luo, S., Zhang, Y., Hu, Q., Chen, J. J., Li, K. P., Lu, C., et al. (2012). Dynamic nucleotide-binding site and leucine-rich repeat-encoding genes in the grass family. *Plant Physiol.* 159, 197–210. doi:10.1104/pp.111.192062
- Ma, L., Yu, Y., Li, C. Q., Wang, P. T., Liu, K. Q., Ma, W. J., et al. (2022). Genome-wide association study identifies a rice panicle blast resistance gene Pb3 encoding NLR protein. *Int. J. Mol. Sci.* 23, 14032. doi:10.3390/ijms232214032
- Minh, B. Q., Nguyen, M. A. T., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi:10.1093/molbev/mst024
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi:10.1093/molbev/msaa015
- Pan, Q. L., Wendel, J., and Fluhr, R. (2000). Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J. Mol. Evol.* 50, 203–213. doi:10.1007/s002399910023
- Parker, J. E., Coleman, M. J., Szabo, V., Frost, L. N., Schmidt, R., vanderBiezen, E. A., et al. (1997). The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *Plant Cell.* 9, 879–894. doi:10.1105/tpc.9.6.879
- 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-Genes Genomes Genet.* 7, 1577–1585. doi:10.1534/g3.117.040485
- Santos, M. D., de Resende, M. L. V., Alves, G. S. C., Huguet-Tapia, J. C., Junior, M., and Brawner, J. T. (2022). Genome-wide identification, characterization, and comparative analysis of NLR resistance genes in *Coffea* spp. *Front. Plant Sci.* 13, 868581. doi:10.3389/fpls.2022.868581
- 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, 9–12. doi:10.1016/j.tplants.2018.10.015
- Shao, Z. Q., Xue, J. Y., Wu, P., Zhang, Y. M., Wu, Y., Hang, Y. Y., et al. (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiol.* 170, 2095–2109. doi:10.1104/pp.15.01487
- 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, 217–234. doi:10.1104/pp.114.243626
- Song, X. M., Wang, J. P., Li, N., Yu, J. G., Meng, F. B., Wei, C. D., et al. (2020). Deciphering the high-quality genome sequence of coriander that causes controversial feelings. *Plant Biotechnol. J.* 18 (6), 1444–1456. doi:10.1111/pbi.13310
- Stolzer, M., Lai, H., Xu, M. L., Sathaye, D., Vernot, B., and Durand, D. (2012). Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. *Bioinformatics* 28, I409–I415. doi:10.1093/bioinformatics/bts386
- Sun, B. J., Wu, L. L., Wu, Y., Zhang, C. F., Qin, L. L., Hayashi, M., et al. (2020). Therapeutic potential of *Centella asiatica* and its triterpenes: A review. *Front. Pharmacol.* 11, 568032. doi:10.3389/fphar.2020.568032
- Walkowiak, S., Gao, L. L., Monat, C., Haber, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi:10.1038/s41586-020-2961-x
- Wan, H. J., Yuan, W., Bo, K. L., Shen, J., Pang, X., and Chen, J. F. (2013). Genome-wide analysis of NBS-encoding disease resistance genes in *Cucumis sativus* and phylogenetic study of NBS-encoding genes in Cucurbitaceae crops. *Bmc Genomics* 14, 109. doi:10.1186/1471-2164-14-109
- Wang, W., Feng, B. M., Zhou, J. M., and Tang, D. Z. (2020). Plant immune signaling: Advancing on two frontiers. *J. Integr. Plant Biol.* 62, 2–24. doi:10.1111/jipb.12898
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49. doi:10.1093/nar/gkr1293
- Whitham, S., Dinesh Kumar, S. P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic-virus resistance gene-n - similarity to toll and the interleukin-1 receptor. *Cell.* 78, 1101–1115. doi:10.1016/0092-8674(94)90283-6
- Xue, J. Y., Zhao, T., Liu, Y., Zhang, Y. X., Zhang, G. Q., Chen, H. F., et al. (2020). Genome-wide analysis of the nucleotide binding site leucine-rich repeat genes of four orchids revealed extremely low numbers of disease resistance genes. *Front. Genet.* 10, 1286. doi:10.3389/fgene.2019.01286
- 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, 165–177. doi:10.1111/jipb.12365
- Zhou, J. M., and Zhang, Y. L. (2020). Plant immunity: Danger perception and signaling. *Cell.* 181, 978–989. doi:10.1016/j.cell.2020.04.028