Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

*Article*

# RNA Interference Pathways in *Juglandaceae*: Based on Gene Diversity, Evolution, and Functional Differentiation

**Yuanpeng Fang[1], Wei Jianming[1], Huang Xin[1], Yunzhou Li [1*] and Xuejun Pan[1,2*]**

[1] College of Agriculture, Guizhou University, Guiyang 550025, PR China
[2] Guizhou Fruit Tree Engineering Technology Research Center, College of Agriculture, Guizhou University, Guiyang 550025, PR China
*Corresponding author: liyunzhou2007@126.com (Y. L.); pxjun2050@aliyun.com (X. P.)

**Abstract:** RNA interference (RNAi) is one of the main mechanisms for disease resistance and small RNA production in plants. The main proteins involved in RNAi include Dicer-like (DCL), RNA-dependent RNA polymerase (RDR), double-stranded RNA-binding (DRB), and Argonaute (AGO). Juglandaceae contains a variety of important woody plants, and walnuts are one of the four major woody plant groups and one of the four major dried fruits in the world. To clarify the evolution and functional differentiation of RNAi-related proteins in the walnut (*Juglans regia*) genome, this study integrated various web resources from gene family acquisition to structural analysis and transcriptome data to correlate walnuts and their congeners. The walnut genome has 5 *DCL*, 13 *RDR*, 15 *DRB* and 15 *AGO* genes, similar genes encoding conserved protein structural domains and conserved motifs with similar subcellular localization. Walnut AGO proteins are classified into three classes and seven subclasses. The DCL is divided into four categories, while RDR is mainly divided into four categories, and DRB can be divided into six categories. The exception is that the copy number of walnut *RDR1* is 9, in which seven *RDR1* are distributed in clusters on chromosome 16. Purifying selection drove the formation of walnut genes, but protein classes were subjected to varying degrees of purifying selection. Additionally, these results showed some similarity in other plants of the walnut family. Moreover, different RNAi-related genes of walnut produced abundant selective expression in response to different tissues and stresses. In this study, *DCL*, *RDR*, *DRB* and *AGO* gene families were identified and analysed in the genome of the walnut family for the first time and preliminarily examined the evolution, structure and expression characteristics of these families to provide a preliminary basis for the evolution of the walnut RNAi pathway and breeding research.

**Keywords:** walnut; RNA interference; Argonaute (AGO); Dicer-like (DCL); RNA-dependent RNA polymerase (RDR); double-stranded RNA-binding (DRB); evolution; expression characteristics

## 1. Introduction

The gene interference (RNAi) pathway is the main pathway of biological resistance to viral hazards and the central pathway of small RNA formation (Baksa & Szittya, 2017; Qin et al., 2018). Existing studies have shown that RNAi is involved in coping with various abiotic and biotic hazard coordination processes and has an important role in plant growth and development (Das et al., 2020, Tang et al., 2017; Sabbione et al., 2019; Mochizuki & Gorovsky, 2005). Dicer-like (DCL), RNA-dependent RNA polymerase (RDR) and Argonaute (AGO) are three classes of proteins that are core components of RNAi. DCL is a protein that enables the conversion of dsRNA to primary siRNA, and secondary siRDR can facilitate the production of more siRNA by retranscribing primary siRNA to synthesize dsRNA, while AGO forms an RNA silencing complex (RIGC) with siRNA, which in turn is involved in the resistance response through RISC. RISC is involved in the resistance response (Fang et al., 2020). Second, double-stranded RNA-binding (DRB) can participate in dsRNA cleavage instead of DCL and contribute to AGO protein binding and downstream small RNA formation; therefore, DRB can also be considered an important

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

component in the RNAi mechanism (Montavon et al., 2017). However, the relatively complex structures of the AGO, DCL, DRB, and RDR proteomes and the abundance of gene copy phenomena have led to difficulties in obtaining homologous genes directly for research. For this reason, RNAi core proteins have been identified in several species, such as hairy poplar (*Populus trichocarpa*), grape (*Vitis vinifera*), citrus (*Citrus sinensis*), tomato (*Solanum lycopersicum*), and rice (*Oryza sativa*) (Bai et al., 2012; Zhao et al., 2015b; Zhao et al., 2015a; Sabbione et al., 2019). In general, DCL proteins (DEAD box, RNA decapping enzyme structural domain, DUF283, ds RBD, RNase III and PAZ) and AGO (including at least one each of PAZ and PIWI structures) proteins are multiconserved structural class proteins, while RDRase has only one RdRP structural domain. PAZ and PIWI structures are associated with the formation of RNA binding pockets, while the ds RBD is a key region for DCL and RNA recognition and binding (Hartman et al., 2013; Zhang et al., 2015). In contrast, DRB proteins include one or more DSRM (double-stranded RNA binding motif) structures, a result related to their hydrolytic activity on dsRNA (Montavon et al., 2017). The expression of DCL, RDR, DRB and AGO genes in a variety of plants is affected by a variety of biotic or abiotic adversities that enhance their expression levels, such as virus infestation, hormonal stress, drought, and low temperature (Boland et al., 2011; Bai et al., 2012; Gan et al., 2017, Esposito et al., 2018). In addition, multiple basic life activities are also associated with *DCL*, *RDR*, *DRB* and *AGO* genes, such as methylation (Zhang et al., 2014), fruit drop (Sabbione et al., 2019), and male sterility (Carbonell & Carrington, 2015).

The eukaryotic RDR family is formed by the expansion of three ancient copies of RDR $\alpha$, RDR $\beta$ and RDR $\gamma$ and further duplication events as major eukaryotes diverge, resulting in the multicopy character of the RDR family (Zong et al., 2009). Dicer followed the evolutionary chain in early plant and animal evolution through independent replication and diversification, continuously reshaping their RNA binding pockets responsible for antiviral immunity through strong selection of proteins. Due to changes in the structure and sequence of its Dicer domain, this led to the specialization of the regulatory or protective functions of animal and plant paralogous homologous genes (Mukherjee et al., 2013). The evolution of the *AGO* gene family antiviral activity may have involved 133-143 duplications and 272-299 loss events, including five major duplications. The divergence of green algae may have formed four major branches of the *AGO* gene family (I: 1/10, II: 5, III: 4/6/8/9, IV: 2/3/7) (Li et al., 2022). Members of the *DRB* gene family have two branches depending on the number of their DSRBs, and the number of DRB proteins varies between species (Clavel et al., 2016). However, the understanding of the overall evolutionary trajectory of the RNAi system is incomplete, and in particular, there is a massive lack of evolution of the RNAi system in woody plants.

Among them, walnuts are known as one of the "four dried fruits", are rich in nutrients, and play an important role in the development of the world's agricultural economy. However, in the walnut cultivation industry, several adversities seriously affect its high quality and quantity production (Wang et al., 2021), such as cold damage, diseases, and insect pests. In recent years, a high-quality walnut genome (*Juglans regia*) has been completely sequenced (Zhang et al., 2020), and the evolutionary study of its RNAi system based on feasible methods is an important part of facilitating the application of model plant research on walnut plants. Second, the process of testing preliminary walnut disease resistance using gene expression will also enhance our understanding of natural resistance to walnut diseases and the functional differentiation of genes (Yan et al., 2019b; Yan et al., 2019a; Feng et al., 2021). To this end, we comprehensively analysed the *DCL*, *RDR*, *DRB*, and *AGO* genes possessed by the walnut genome and investigated their structure, evolution, and possible interactions. Second, the analysis of potential RNAi-related expression signatures in walnut using transcriptomic data will enhance our understanding of natural disease resistance and gene functional differentiation in walnut. To this end, we the DCL, RDR, DRB, and AGO genes were comprehensively analysed that are present in the walnut genome and investigated their structure, evolution, and possible interactions. Data were used to analyse their gene expression differentiation for RNAi-related

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

genes, and these findings are expected to lay the foundation for the study of walnut breeding.

## 2. Materials and Methods

### *2.1. Identification of RNAi-associated protein family genes*

*AGO*, *DCL*, *DRB*, and *RDR* genes were identified using the following: 1. Arabidopsis and tomato homologous proteins by blastp search; 2. HMMER 3.61 was used to search for marker domains (*RDR* "RdRP", *AGO* "PAZ, PIWI", *DCL* "PAZ, dsRBD", *DRB* "DSRM") (Potter et al., 2018; Cui et al., 2020; Sabbione et al., 2019); 3. SMART (http://smart.embl-heidelberg.de/) (Letunic & Bork, 2018) and PFAM (http://pfam.xfam.org/) (El-Gebali et al., 2019) websites were used to determine the marker structure domain completeness of the marker structure. Finally, manual merging and deduplication were performed.

### *2.2. Conservative and chromosomal localization analysis*

Conservative motifs, gene structure maps and chromosome localization maps were predicted or drawn using MEME (search motif number 10 for each family, length limit 6-100) and TBtools v.1082 software (Bailey et al., 2009; Chen et al., 2020). In addition, the walnut genome used in this paper was obtained from the Hardwood Genome Database (https://hardwoodgenomics.org/), Arabidopsis from TAIR (www.arabidopsis.org/), tomato genome from Solanaceae genome (https://solgenomics.net/), and other genomes from EnsembelPlant (http://plants.ensembl.org/). Other genomes not mentioned in the methods were obtained from 3 databases (http://gigadb.org/, https://ngdc.cncb.ac.cn/, http://www.juglandaceae.net/).

### *2.3. Prediction of base information and interactions*

Protein isoelectric points and relative molecular masses were predicted using EXPASy's Prot-Paramtool online program (http://web.expasy.org/cgi-bin/protparam/protparam) (Duvaud et al., 2021); protein subcellular localization was predicted using Softberry (http://www.softberry.com/). AGO, RDR, DRB, and DCL interactions between proteins were predicted using STRING (https://string-db.org/) and plant.map (http://plants.proteincomplexes.org/) websites (species mainly referenced to Arabidopsis, with a screening restriction of score >0.4) (Szklarczyk et al., 2017; McWhite et al., 2020) and finally by Cytoscape 3.71 (Shannon et al., 2003) for presentation.

### *2.4. Evolutionary analysis*

Tree building analysis was calculated using MEGA 7.0 based on AGO, DCL, DRB and RDR protein sequences (Clustal W method alignment, posterior value bootstrap=1000, and neighbour-joining method for tree building) (Das et al., 2020; Kumar et al., 2016). The evolutionary analysis based on repetition type was calculated using MCScan X (Wang et al., 2012). The screening principles for tandem duplicate gene pairs are based on (1) the length of coverage on the pair exceeds 75% and the homology is 75%, and (2) the gene distribution should be close in physical distribution (Gu et al., 2002). Ka/Ks (synonymous substitution/nonsynonymous substitution) was calculated using TBtools v.1082, and differentiation time was calculated using T=Ks/2r ($r=1.5*10^{-8}$) (Chen et al., 2020; Huang et al., 2015). Images were plotted using TBtools, R v.4.02, and Evolview (https://www.evolgenius.info/evolview/) (Subramanian et al., 2019) tools.

### *2.5. Analysis of promoters and the expression of RNAi-associated genes in walnut*

To explore the expression levels of RNAi-associated genes in different walnut tissues and under different stresses, raw sequence data (SRP034866, SRP290361, SRP359475, SRP343997) was used in fastq format from SRA (https://www.ncbi.nlm.nih.gov/sra) (Fang et al., 2021a). High-quality sequenced samples were then compared to the Hardwood Genome Database walnut genome using HISAT2 (Kim et al., 2019). For unique mapped reads, raw counts were aggregated using featureCount (Liao et al., 2014) and a custom

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

Perl script, and transcript per million (TPM) values were calculated for each gene. The sorghum upstream sequence (1,000 bp before the 5th end of the coding sequence) was selected as the promoter sequence and submitted to the PlantCare website (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)(Lescot et al., 2002), and TBtools was used to visualize the results.

## 3. Results and Analysis

### *3.1. Identification of RNAi-associated genes*

After careful analysis and identification, it was found that walnut has 15 *AGO* genes, 5 *DCL* genes, 15 *DRB* genes, and 13 *RDR* genes. The lengths of the AGO and DCL proteins ranged from 900-1100 aa and 1400-2000 aa, respectively, while the lengths of the RDR proteins, except for JrRDR1e (386 aa) and JrRDR1f (630 aa), were approximately 1000 bp. While the DRB protein sizes are mostly distributed at 400-600 aa, some of them are near 200 aa (JrDRB6.1, JrDRB6.2, JrDRB7.2, JrDRB7.3) or near 1000 aa (JrDRB7.1).

In addition, the molecular mass of most of this walnut AGO, RDR, and DCL proteins was 100-200 kDa, with the largest protein being DCL1 (224 kDa) and the smallest being JrRDR1e (44 kDa). In contrast, the molecular weights of walnut DRB proteins are mainly located at 40-60 kDa, with the smallest protein being much smaller than other RNA-related proteins (JrDRB7.3, 18 kDa). The isoelectric point of the walnut AGO protein was between 8 and 9, and the subcellular localization was in the cytoplasm or nucleus, except for JrAGO2a, which was in the extracellular (secretory) region. The isoelectric point of the walnut DCL protein was between 6 and 8, and the subcellular localization was in the nucleus, while the isoelectric point of the walnut RDR protein was larger at between 5 and 9, and the subcellular localization was in the extracellular (secretory) region. The isoelectric points of walnut DRB proteins are mainly between 8 and 10, with a few between 5 and 6 (JrDRB4A, JrDRB7.2, JrDRB7.1), and the subcellular localization is dominated by the nucleus and chloroplasts. Interestingly, the subcellular localization of JrDRB7.4 showed the highest probability of being in the vesicles (Supplementary Table 6).

### *3.2. Phylogenetic tree analysis of RNA-associated proteins*

Figure 1 shows the phylogenetic relationships of walnut with rice, poplar mullein, citrus, grape, tomato, and Arabidopsis AGO, RDR, DRB, and DCL proteins. To characterize the more likely true protein homology relationships, phylogenetic tree positions were used to name the walnut genes. As shown in Figure 2A, AGO is mainly divided into three classes (I, II, and III), and walnut AGO proteins are distributed in seven subclasses among these three classes (I: AGO1, AGO5, AGO10; II: AGO2/3, AGO7; III: AGO6, AGO4). The copy number of AGO1, AGO4, and AGO10 proteins is 3, while the copy number of AGO2 and AGO5 proteins is 2, and the copy number of AGO6 and AGO7 proteins is 1. In addition, walnut AGO proteins have the highest homology with grape or citrus AGO proteins among most members. However, specifically, the three proteins of walnut AGO10, JrAGO10a and JrAGO10c formed a direct homologous pair with CsAGO10, while JrAGO10b formed a direct homologous pair with VvAGO10b. Walnut JrAGO6 formed a direct homologous pair with PtAGO16 but not with grape or citrus AGO proteins.

Figure 1B shows that *DCLs* are divided into four branches, *DCL1*, *DCL2*, *DCL3*, and *DCL4*, and walnut contains 1, 2, 1, and 1 gene in each of the four types of DCLs. jrDCL2s and jrDCL4 have the highest homology with citrus, followed by grape. JrDCL1 and JrDCL3 had the highest homology with grape, followed by citrus and mullein. Figure 1C shows that *RDR* is mainly divided into four branches, *RDR1*, *RDR2*, *RDR3*, and *RDR6*. Walnut *RDR2* and *RDR3* are single copies, while the RDR6 copy number is 2, and the RDR1 copy number is 9. Walnut RDR3 has the highest homology with grape RDR3, while walnut RDR1, RDR2, and RDR6 have the highest homology with citrus or burkegonia. Figure 1D shows that *DRB* can be clustered in the six published DRB branches (*DRB1*, *DRB2*, *DRB3-5*, *DRB4*, *DRB6*, and *DRB7*) and that *DRBs* are multicopy in all classes except *DRB4*, which is a single copy and has copy numbers of 2, 2, 4, 2, and 4, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

Walnut and *P. trichocarpa* and *C. sinensis* had the highest homology, and proteins from these three species tended to cluster on the same branch. In addition, in some categories, walnut DRB proteins showed a clear genetic gap between two (e.g., JrDRB7.1, JrDRB7.2, and JrDRB7.3 clustered on one node each and did not cluster together).

### *3.3. Basic structural analysis of RNA-associated genes in walnut*

The analysis of walnut genes and protein structures revealed that 15 *AGO* genes have 7 structural domains, 5 of which are shared, and the rest of ArgoMid is distributed in the class I proteins of walnut AGO. Gly-rich_Ago1 is only distributed in walnut AGO1 proteins; 5 *DCL* genes have 9 structural domains, 7 of which are shared, and dsrm is only distributed in JrDCL1 and JrDCL4 proteins; 13 RDR proteins have a segment of RdRP, and DND1_DSRM is distributed in JrDCL1 and JrDCL4 proteins. In addition, the dsrm domain is only distributed in the JrDCL1 protein, and the DND1_DSRM domain is distributed in the JrDCL1 and JrDCL4 proteins; all 13 RDR proteins possess a segment of the RdRP structural domain, which is at the front end of the protein sequence. DRB is mainly dominated by the distribution of different numbers of DSRM structures, where 66% of the nucleolar DRB possesses a typical DSRM structural domain distribution. Then, 27% of the DRB proteins possess only 1 DSRM structure at the C-terminus, while JrDRB4A possesses 3 DSRM structural domains. It is particularly interesting that the JrDRB7.1 protein, which possesses only 1 DSRM, also possesses 1 Methyltransf_11 structure (Table 1).

The gene structure showed that the number of introns of class III genes among the three classes of *AGO* was only 3, while the intron structure of class I and II genes was complex (>20 introns) (Figure 2A 1&3). Five *DCL* genes all had more than 20 introns (Figure 2B 1&3). While the basic intron number of the *RDR* gene was less than 6, the *JrRDR3* gene possessed 19 introns (Figure 2C 1&3). The walnut *DRB* gene as a whole showed a simple intron–exon structure, with all *DRBs* except *JrDRB7.1* (8 introns) possessing only 4 or fewer introns (Figure 2D 1&3).

Ten conserved motifs were retrieved in each family, named AGO-1~10, DCL-1~10, RDR-1~10, and DRB-1~10. Parts 1 & 2 of Figure 2 show that the conserved motifs of the DCL and AGO families are widely distributed in each walnut protein, but the distribution of AGO-1 (class I proteins and AGO6) and DCL-6 (DCL2s and DCL4) is limited (Figure 2A and B 1 & 2). While the walnut RDR proteins are slightly less conserved, they possess at least RDR-1 and RDR-2; the remaining structures are present in 2 unequal proteins, and notably, RDR-8 is present in most of the RDR1s and RDR2 proteins. Because the DRB proteins are more specific, all motifs except DRB-3 were missing in different classes (Figure 2C 1&2). Among them, DRB3-5 is characterized by possessing intrinsic DRB-4, while JrDRB4A presents the structure of both ends of DRB-1 (Figure 2D 1&2). The motifs of AGO-1, DCL-6, DRB-3, and RDR-8 proteins were briefly extracted and found that a large number of them were extremely conserved (Figure 2E).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

**Table 1.** The domains of AGO, DCL, DRB and RDR proteins in walnut.

| Protein name | Domain 1 | Domain 2 | Domain 3 | Domain 4 | Domain 5 | Domain 6 | Domain 7 | Domain 8 | Domain 9 |
|---|---|---|---|---|---|---|---|---|---|
| DCL | Ribonuclease_3 | Ribonucleas_3_3 | Dicer_dimer | PAZ | Helicase_C | dsrm | DND1_DSRM | ResIII | DEAD |
| JrDCL1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| JrDCL2a | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| JrDCL2b | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| JrDCL3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| JrDCL4 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| AGO | Piwi | PAZ | ArgoN | ArgoL1 | ArgoL2 | ArgoMid | Gly-rich_Ago1 | | |
| JrAGO10b | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | |
| JrAGO10a | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | |
| JrAGO10c | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | |
| JrAGO1a | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| JrAGO1b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| JrAGO1c | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| JrAGO2a | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO2b | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO4a | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO4b | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO4c | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO5a | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | |
| JrAGO5b | 1 | 1 | 1 | 1 | 1 | 1 | 0 | | |
| JrAGO6 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | | |
| JrAGO7 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | | |
| DRB | DSRM | | | | | DSRM+ Methyltransf_11 | | | |
| 1 | JrDRB6.1 JrDRB6.2 JrDRB7.2 | | | | | JrDRB7.1 | | | |
| 2 | JrDRB1.1 JrDRB1.2 JrDRB2.1 JrDRB2.2 JrDRB3-5.1 JrDRB3-5.2 JrDRB3-5.3 JrDRB3-5.4 JrDRB7.3 JrDRB7.4 | | | | | | | | |
| 3 | JrDRB4A | | | | | | | | |
| RDR | RdRp | | | | | | | | |
| JrRDR1a JrRDR1b JrRDR1c JrRDR1d JrRDR1e JrRDR1f JrRDR1g JrRDR1h JrRDR1i JrRDR2 JrRDR3 JrRDR6a JrRDR6b | | | | | | | | | |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

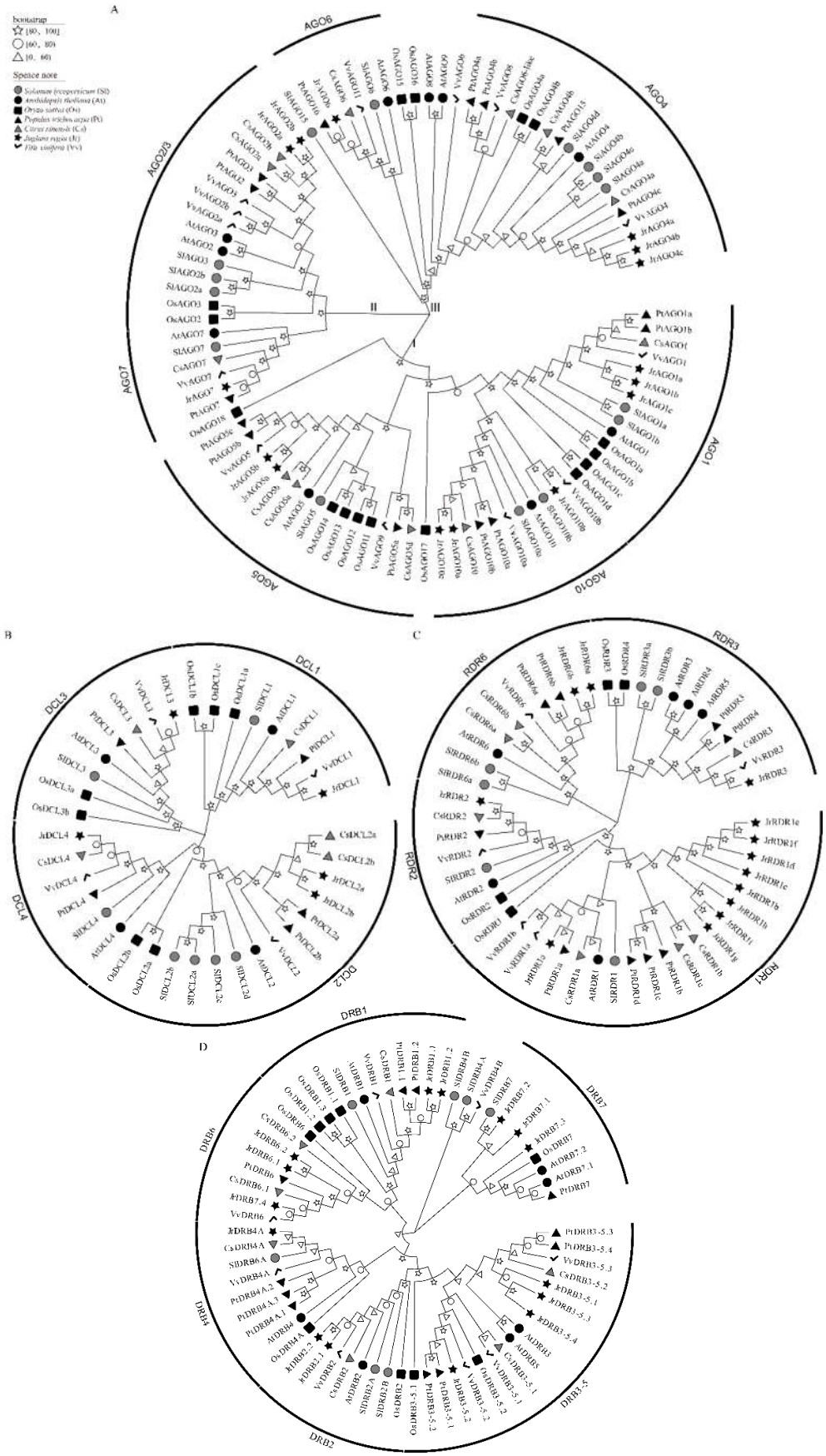


**Figure 1. Phylogenetic tree analysis of AGO, RDR, DRB and DCL proteins**.

Note: Oryza sativa (Os); Arabidopsis thaliana (At); Solanum lycopersicum (Sl); Vitis vinifera (Vv); Populus trichocarpa (Pt); Citrus sinensis (Cs); Juglans regia (Jr).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

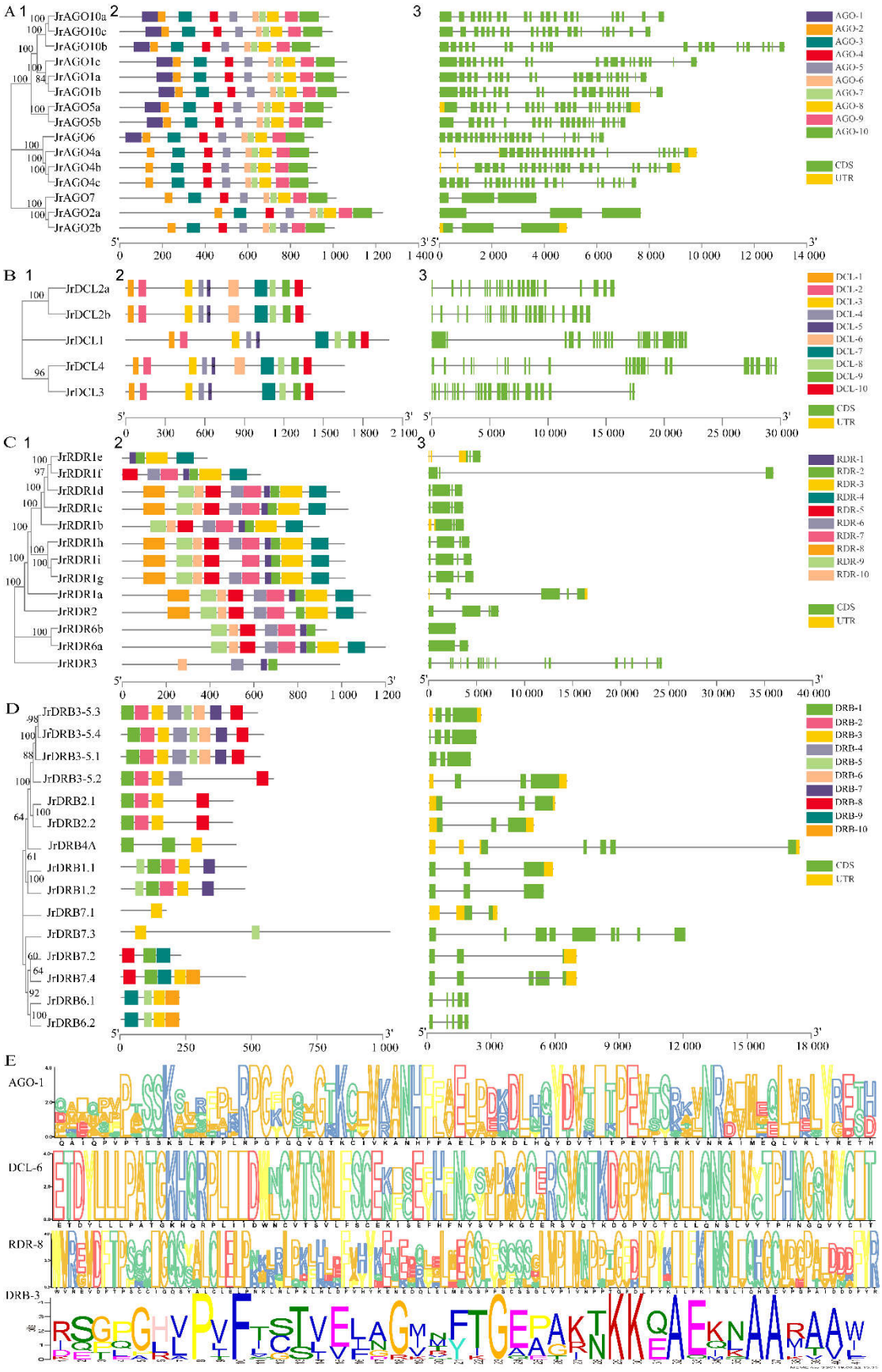


**Figure 2.** Phylogenetic analysis (1), protein motif (2) and gene structure (3) of walnut AGO, RDR, DRB and DCL.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

*3.4. Evolutionary analysis of RNA-associated genes in walnut*

To investigate the evolutionary history of the *AGO*, *RDR*, and *DCL* genes in walnut, collinearity analysis was performed. Figure 3A shows that five *DCL* genes in walnut were distributed on chromosomes 5, 6, 7, 8, and 11. Thirteen *RDR* genes were distributed on chromosomes 1, 2, 3, 11, and 16, with the exception of RDR1b (Jr01) and RDR1c (Jr02), and nine clusters of *RDR* genes were arranged in the middle and upper parts of chromosome 16. Fifteen *AGO* genes were distributed on chromosomes 1, 3, 4, 5, 6, 7, 10, 12, 13, 14, and 15, forming five collinearity pairs. The 14 *DRB* genes were distributed on chromosomes 1, 3, 4, 5, 6, 8, 9, 10, 11, and 14, while *JrDRB6.2* was located on the fragment that could not be loaded onto the chromosome. Two *DCL2* genes, 4 *DRB* genes, and 9 *AGO* genes formed duplicated gene pairs resulting from the duplication of the fragment. In addition, all these collinearity pairs generated by fragment duplication were formed by strong purifying selection (Figure 3A and 4).

Colineages were relatively abundant in woody plants (23 pairs in grapes, 24 pairs in citrus, and 32 pairs in burgundy poplar). Grapes are a vine woody plant that formed 23 collinearity pairs with 23 genes of walnuts (including 9, 3, 7, and 4 collinearity pairs of *AGO*, *DCL*, *DRB*, and *RDR* genes); notably, among them, collinearity pairs of *DRB4*, *DRB6*, *DCL3*, and *AGO1* were not found at this stage. Citrus aurantium is a domesticated and important woody fruit tree that forms 24 collinearity pairs with 24 walnut genes (including 11, 2, 7, and 4 collinearity pairs for *AGO*, *DCL*, *DRB*, and *RDR* genes). Burkholderia is an old woody plant that forms a total of 32 collinearity pairs with 20 walnut genes (including 13, 3, 10, and 6 collinearity pairs for *AGO*, *DCL*, *DRB*, and *RDR* genes); interestingly, collinearity pairs for *AGO1* are accessible in citrus and burkholderia, while collinearity pairs for *DRB4*, *DRB6*, and *DCL3* are similarly lacking, and covariance with walnut DCL4 could not be found in citrus and woolly poplar (Figure 3B).

Only three pairs of collinearity (involving three genes, *JrRDR1d*, *JrAGO10b*, *JrDRB2.2*) were found between monocotyledonous rice and dicotyledonous walnut; relatively few collinearity were also found between herbaceous dicotyledonous plants and walnut; 18 pairs of collinearity involved 15 walnut genes (involving *AGO10*, *AGO1*, *AGO4*, *AGO5*, *AGO6*, *JrAGO7*, *DRB2*, *DRB3-5*, *DRB7*, *RDR1*, *RDR3*) in Arabidopsis; and 22 pairs of collinearity involved 19 walnut genes (involving *AGO10a*, *AGO1*, *AGO2*, *AGO4*, *AGO5*, *AGO6*, *AGO7*, *DCL1*, *DCL2*, *DRB1*, *DRB2*, *DRB7*, *RDR3*) in tomato. (Figure 3C).

In addition, all collinearity pairs between species were formed by purifying selection, but the degree of purifying selection differed among walnut *DCL*, *AGO*, *DRB*, and *RDR* genes. Little difference was noted in the degree of purifying selection and divergence times among the six species, but the internal collinearity pairs of walnut were subject to weaker purifying selection pressure (~2-fold weaker) and diverged much later (~10 Myr) than the divergence of other genes among the species (30-70 Myr) (Figure 4A & B). *RDR1* (~20 Myr) diverged later than *RDR3* (~45 Myr). *DCL2s* (~0.27) were subjected to much less purifying selection pressure than *DCL1* (~0.14) genes, while *DRB2* (~0.139) was subjected to much more purifying selection pressure than *DRB3-5* (~0.257) and *DRB7* (~0.307) genes and much higher selection pressure for purification (Figure 4C, D & E). For the *AGO* gene family, *AGO-I* (~0.079) was subjected to much higher purification selection pressure than the *AGO-II* class (~0.205); specifically, *AGO* genes *AGO2s* were subjected to much lower purification selection pressure than *AGO4s* and *AGO7s*, while *AGO5s* were subjected to much lower purification selection pressure than *AGO10s* and diverged earlier (Figure 4F, G & H).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

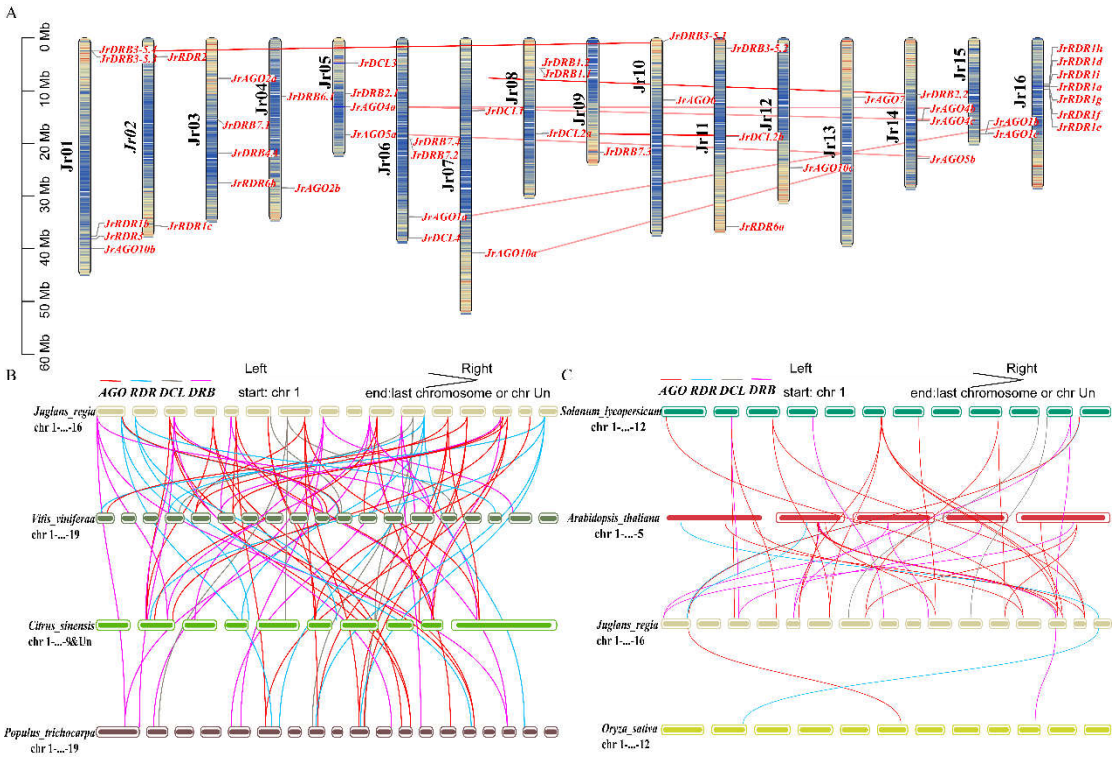


**Figure 3.** Chromosomal localization and intraspecific collinearity of *AGO*, *RDR*, *DRB* and *DCL* in walnut (A), collinearity of *AGO*, *RDR*, *DRB* and *DCL* between woody plants and walnut (B), and collinearity of *AGO*, *RDR*, *DRB* and *DCL* between herbaceous plants and walnut (C).

Note: The colour depth of the chromosome in Figure A represents the gene distribution density within 10000 bp. The darker the colour is, the higher the gene density.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

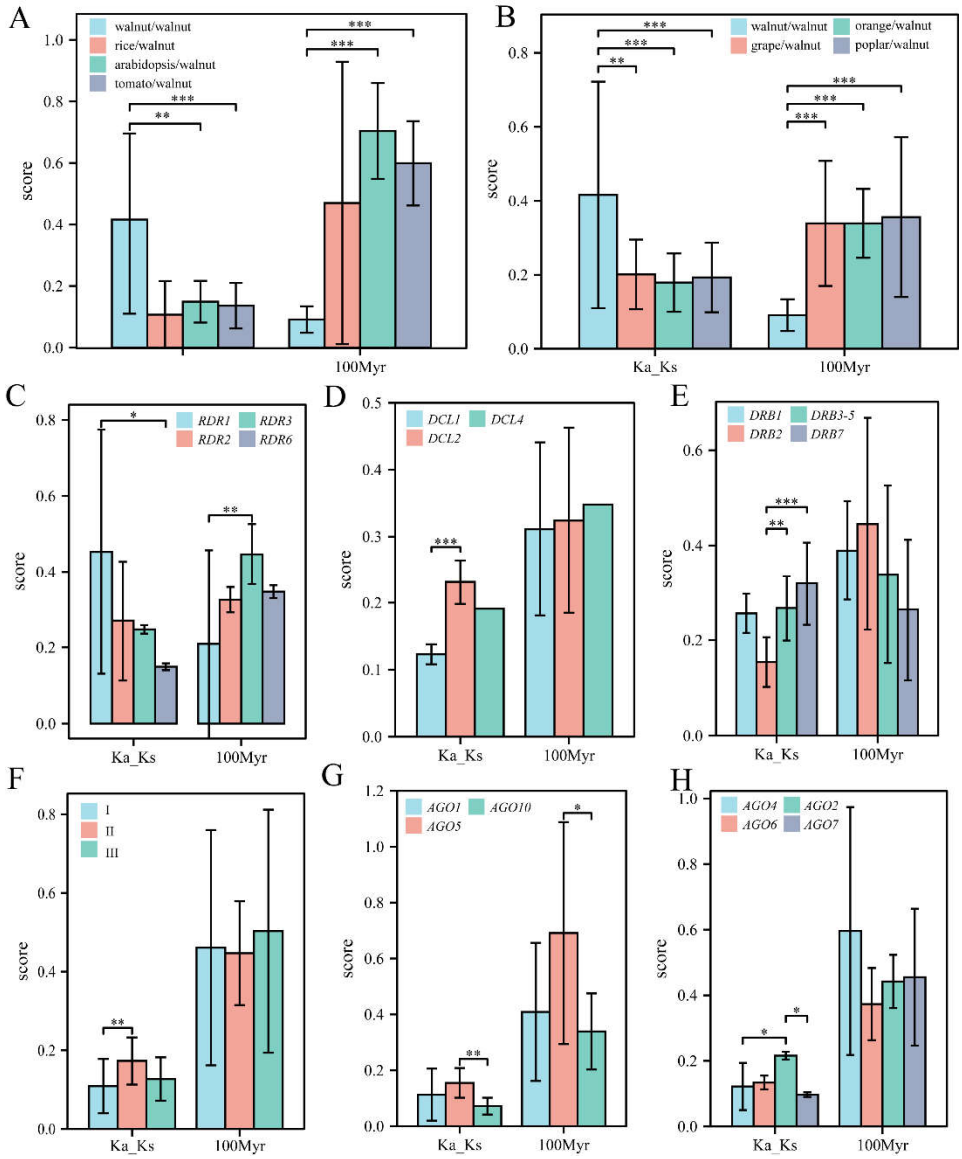


**Figure 4.** Evolution rate (Ks/Ks) and differentiation time analysis of collinear gene pairs between *AGO*, *RDR*, *DRB*, *DCL* and different species in walnut. Evolution rate (Ks/Ks) and differentiation time analysis in (A) non-woody plants, (B) non-woody plants, (C) different *RDR*, (D) different *DCL*, (E) different *DRB*, (F) different *AGO* type, (G) different *AGO*-I type protein, (H) different *AGO*-II and III type protein.

Note：The significance marker was NS, P ≥0.05; *, p < 0.05; * *, p < 0.01; * * *, p < 0.001.

### *3.5. Evolutionary analysis of RNA-associated genes in different dried fruits*

Figure 5 shows that tandem repeats have an important role in the copy duplication of RNAi-related families in walnut, where it wasfound 21 possessing possible tandem repeat relationships, which consisted of 1 *JrRBD7.4-JrRBD7.2* relationship and 2 *AGO1*, *AGO4*, *DRB1*, *DRB3-5* and 6 *RDR1* (Figure 5A). Moreover, the relationship is composed of genes on 6 chromosomes, independent of the other 10 chromosomes. It is particularly interesting that all 2 AGO genes of chromosome 15 form tandem repeats first. Further analysis revealed that these gene pairs included 3 pairs of genes with very short evolutionary times (Ka/Ks=0, *JrDRB3-5.4-JrDRB3-5.3*, *JrDRB1.2-JrDRB1.1*, *JrAGO4b-JrAGO4c*) and 1 pair of positively selected genes (*JrRDR1f- JrRDR1e*), and the selection pressure of the remaining genes was purified selection. In addition, the distribution time of *AGO1* is approximately 7.6 Myr (*JrAGO1b-JrAGO1c*), while the differentiation time of *DRB7* is approximately 0.9 Myr (*JrDRB7.4-JrDRB7.2*) and that of *RDR1* is approximately 1.5-11 Myr.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

Uniquely, the divergence time of the *JrRDR1f- JrRDR1e* gene pair subjected to positive selection pressure was extremely close at only approximately 0.13 Myr (Figures 5B and 5C).

Figure 6 shows the collinearity between walnut RNAi-related families and two other important dried fruits, and 24 pairs of collinearity were found between walnut and two different dried fruits. The collinearity between walnut and filbert involved 22 walnut genes, while the collinearity with apricot involved 24 different walnut genes. Interestingly, 17 genes could be detected in both collinearities, including *AGO*, *DCL*, *DRB*, and *RDR* genes in the numbers 9, 1, 4, and 3 (Figure 6A). After careful examination, it was found that the genomes of apricot and walnut had less *AGO2* and *DRB1* and more *DCL2-4* than filbert, while *DRB4* and *DRB6* did not show collinearity in either dried fruit. In addition, the purifying selection pressure of walnut internal genes was less than the interspecific purifying selection pressure of different dried fruit plants, while the purifying selection pressure of *AGO* genes was also greater than that of *DRB* and *RDR*; additionally, there was no significant difference in selection pressure of single subtle different gene classes (Figure 6B, C & D). In addition, gene pairs within walnut diverged later than those within apricot and even later than those within filbert; however, there was no significant difference in divergence time between gene classes (Figure 6B, C & D).

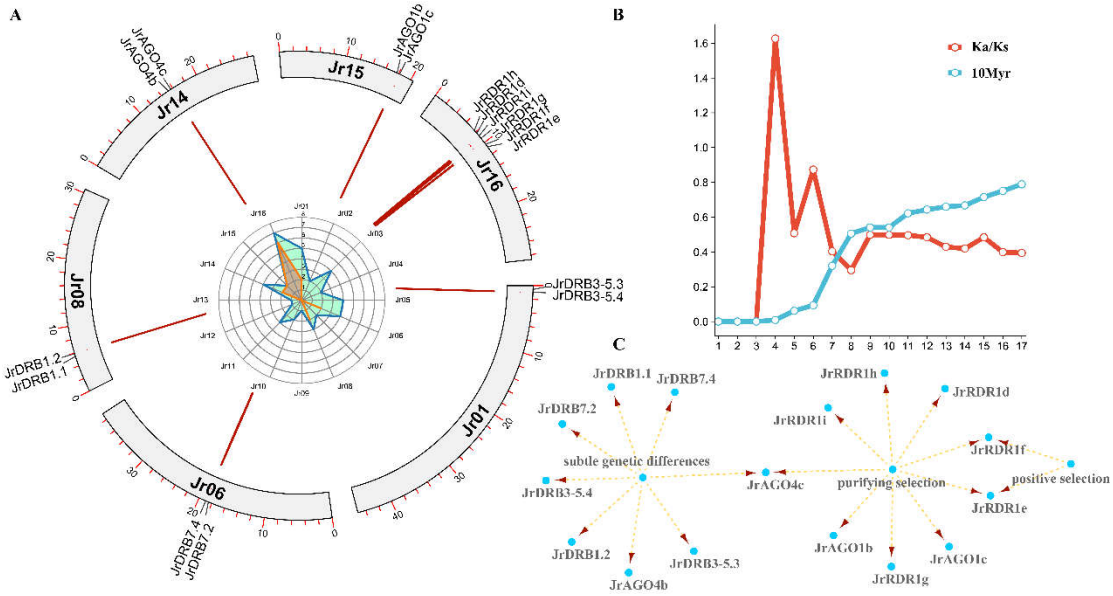


**Figure 5.** Detection of tandem repeats in walnut *AGO*, *RDR*, *DRB* and *DCL*. (A) Distribution characteristics of tandem repeat gene pairs, (B) evolutionary rate (Ks/Ks) and differentiation time distribution, (C) selection pressure analysis of different genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

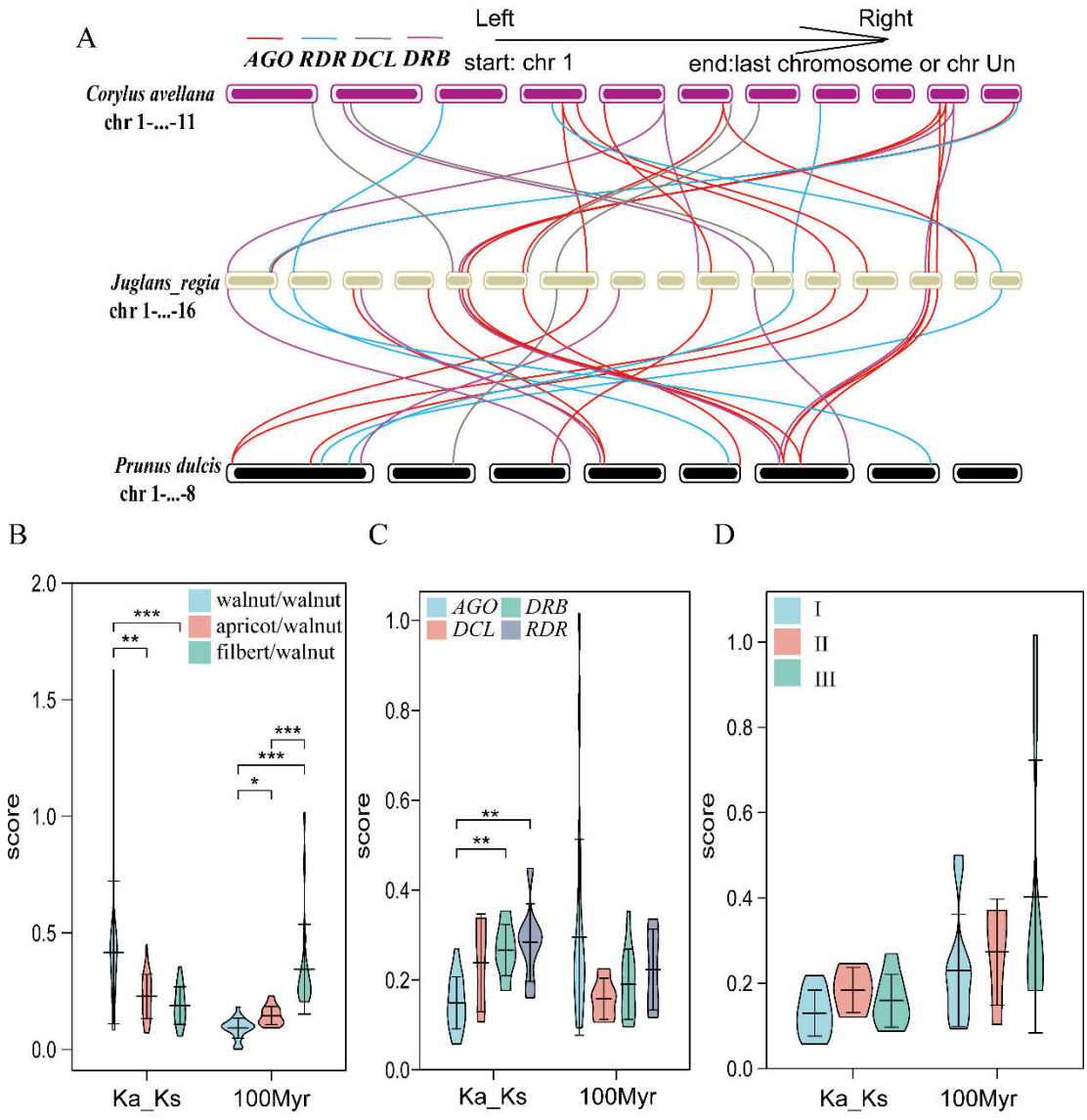


**Figure 6.** Intraspecific collinearity of *AGO*, *RDR*, *DRB* and *DCL* in *Prunus dulcis*, *Corylus avellana* and walnut (A), and evolution rate (Ks/Ks) and differentiation time analysis of (B) different plants, (C) different RNAi gene types, and (D) different *AGO* types.

### *3.6. Evolutionary analysis of RNA-associated proteins in Juglandaceae*

In this section, the potential evolution of RNA-associated proteins was parsed in *Juglandaceae*. It was first analysed the three accessible species of the proximate genus (*Pterocarya stenoptera*, *Carya illinoinensis*, *C. cathayensis*), six congeneric species (*J. mandshurica*, *J. cathayensis*, *J. hindsii*, *J. microcarpa*, *J. nigra*, *J. sigillata*) and one hybrid (*J. microcarpa x regia*) to identify RNA-associated proteins in their genomes. The results showed that we identified 181, 40, 105, and 67 possible coding genes for *RDR*, *DCL*, *AGO*, and *DRB*. Our preliminary analysis of their protein physicochemical properties showed that RDR, DCL, AGO, and DRB of all Juglandaceae maintained large molecular weights overall, with isoelectric points mainly distributed approximately at 5-8 (Additional Table 3).

When these possible proteins were examined, the presence of 2 RdRP domains was noted in a small number of RDR proteins, while the presence of dsrm structural domains in DRB proteins was widely distributed in different numbers (one, two, three, four). In contrast, in DCL and AGO, abundant structural domains are present, while a considerable abundance of structural domains are missing. For example, 68 proteins possess the distribution characteristics of Piwi, PAZ, N-terminal, L1, L2, and Mid structural domains, while in 24 proteins, Mid structural domains are missing (Additional image 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

Furthermore, a phylogenetic tree of each RNAi-related protein class was constructed from 17 species. The results showed that, on the premise that the protein clusters of each subclass were correctly separated, for AGO, RDR, DRB, and DCL proteins, the proteins of the pecan family were individually separated from those of other plants to form a unique branch. Notably, the pecan family appears to be younger than the RNAi-associated proteins of rice, grape, Arabidopsis, and six other species, as their branches are all at the end of each class of proteins rather than between proteins of several species (Figure 7A-D).

To correctly interpret the evolution of RNAi-associated proteins in the peccary family, the quantitative differences among protein subclasses were compared and showed that all regular protein taxa could be found for almost all species (DCL: DCL1, DCL2, DCL3, DCL4; RDR: RDR1, RDR2, RDR3/4/5, RDR6; DRB: DRB1, DRB2, DRB3-5, DRB6, DRB7; AGO: AGO1, AGO2/3, AGO4, AGO5, AGO6, AGO7, AGO10). The genomes of some of these species appear to be amplified for many gene taxa, such as *RDR1*, *AGO5*, *DRB3-5*, *RDR6*, and *DCL2*. More typically, the Walnut family seems to have a consistent expansion for *RDR1*, with a large number of *RDR1s* found in various Walnut species (Figure 7E).

Meanwhile, collinearity analysis in walnut species was performed at the complete chromosome level and measured the selection pressure and divergence time of collinearity pairs. The results showed that there were 42 and 46 collinearity pairs in the genomes of walnut, iron walnut and Mandshurian walnut, which involved 29 and 33 walnut genes, respectively. In addition, 47 collinearity pairs involving 33 walnut genes existed between walnut and hybrid walnut (*J. microcarpa x regia*). A closer examination revealed that there were significant differences in the divergence times between the genes within walnut and the collinearity pairs of *J. microcarpa*, and in addition, there were significant differences in the relative divergence times of the relationships between walnut in *J. microcarpa* and hybrid walnut. In addition, the *RDR* gene had significantly less selection pressure for purification between these different walnut species than the *AGO* gene, but the difference in selection pressure between subclasses was not significant. Differentiation times and selection pressures for all the same protein subclasses with walnuts showed significant differences between species and genes in different classes, where the differences in purification selection pressures between gene classes were *AGO* < *DRB* and *DRB* < *DCL* < *RDR*, and the differences in their differentiation times showed that *DCL* differentiated later than several other RNAi-associated proteins (Supplementary Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

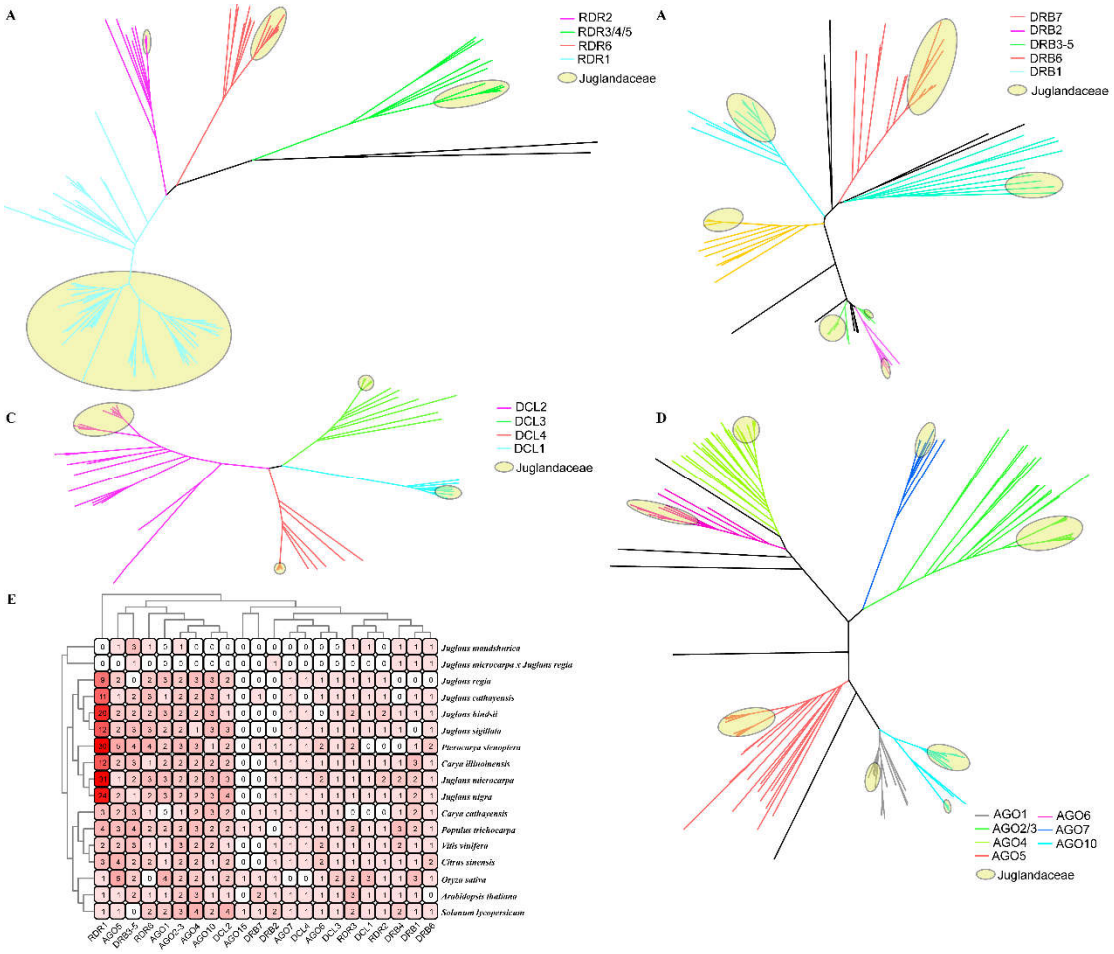


**Figure 7.** Phylogenetic analysis of *RDR* (A), *DRB* (B), *AGO* (C) and *DCL* (D) and quantity distribution (E) in Juglandaceae.

### *3.7. Analysis of walnut RNA-associated protein interactions*

To briefly analyse the function of walnut AGO, RDR, DRB, and DCL proteins, interactions were analysed based on high-confidence coexpression (145) or protein data (38) from Arabidopsis. It was noted an abundance of proteins that may interact with AGO, RDR, DRB, and DCL proteins; these included AGO-DCL-RDR-DRB internal interactions. Among them, the internal interactions include 75 pairs, and these interactions show interactions between AGO and RDR and DCL and DRB; second, DCL proteins can interact with RDR and DRB; in addition, DCL and DRB can have abundant interactions between different classes of proteins (Figure 8).

Furthermore, among these RNAi-associated proteins, only AGO identified high-confidence interacting proteins, and these AGO proteins generated 70 interactions with various other classes of proteins. These include proteins associated with various resistance (asparagine-rich response protein NRP and selenium-binding protein), proteins associated with epigenetic modifications (HTA and HTR), proteins associated with disease resistance (RPRR and MAPKKK class), receptor proteins (LRR), pretranscriptional regulatory elements (arginine-/serine-rich splicing factor RS and translation initiation factor) (SUI), etc. In addition, different proteins of the same class have richer differences in interactions (Figure 8).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

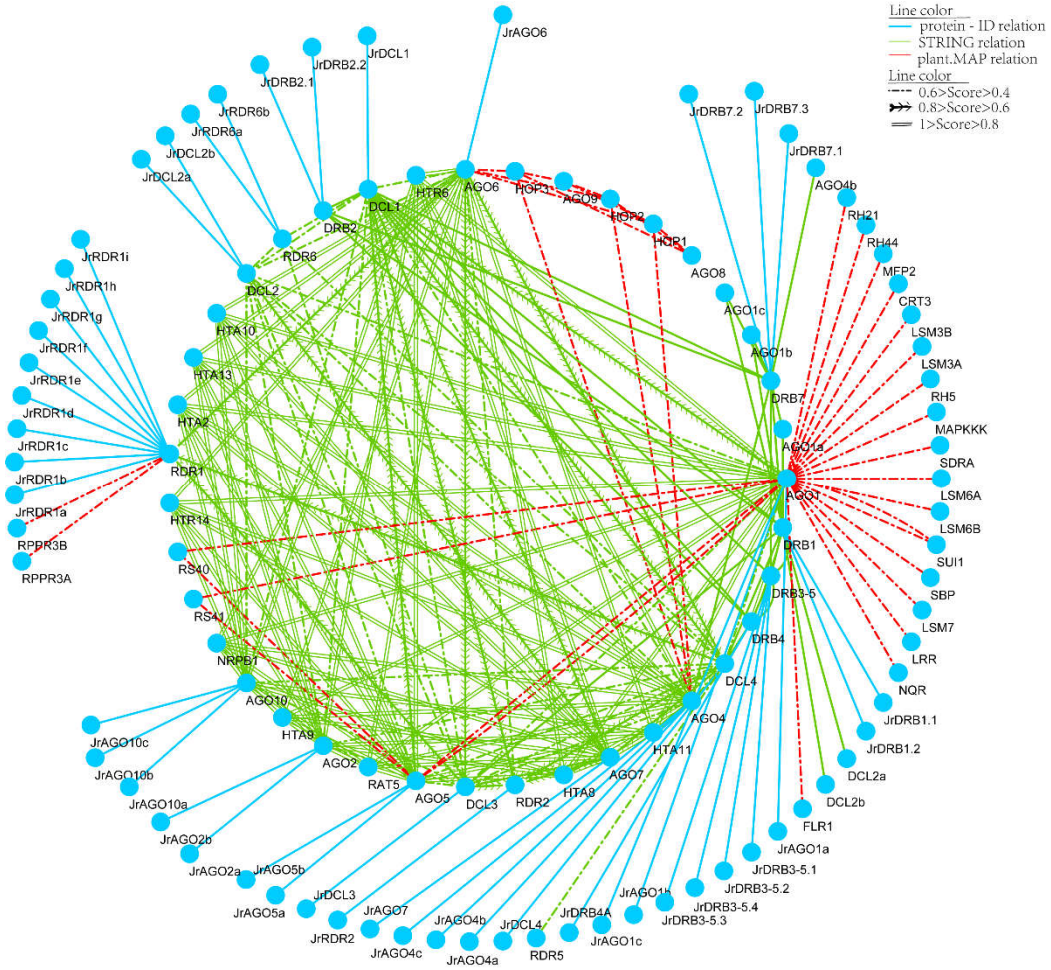


**Figure 8.** Interaction protein analysis of AGO, RDR, DRB and DCL.

Note：FLR, Leucine-Rich Repeat Protein FLOR; HOP, Hsp70-Hsp90 Organizing Protein; HTA, Histone H2axa; HTR, Histone H3-Like; LRR, LRR-Repeat Protein; LSM, Sm-Like Protein; MAPKKK, Mitogen-Activated Protein Kinase Kinase; MFP, Multifunctional Protein; NQR, NADPH: Quinone Oxidoreductase; NRPB, DNA-Directed RNA Polymerase II Subunit; RAT, Resistant to Agrobacterium Transformation; RH, DEA(D/H)-Box RNA Helicase; RPPR, Ribosomal pentatricopeptide repeat protein; RS, Arginine/Serine-Rich Splicing Factor; SBP, Selenium-Binding Proteins; SDRA, Short-Chain Dehydrogenase/Reductase; SUI, Translation Initiation Factor.

### *3.8. Analysis of the expression of RNAi-related genes in walnut*

To explore the functional differentiation of walnut RNAi-associated genes, a preliminary analysis was performed based on promoter elements. The results revealed the presence of abundant cis-regulatory elements of RNAi-related genes, which contained many tissue-specific elements (e.g., seed-specific regulation: RY-element; meristem expression: CAT-box; endosperm expression: GCN4_ motif) and adversity-related elements (e.g., defence and stress responsiveness: TC-rich repeats; low-temperature responsiveness: LTR; maximal elicitor-mediated activation: AT-rich sequence) (Supplementary Figure 3 and Supplementary Table 5). This suggests that walnut RNAi-related genes may be associated with tissue-specific expression and stress response.

Furthermore, the expression was analysed from transcriptome data of walnut RNAi-related genes. The results showed that there were some levels of tissue specificity of RNAi-related genes in different tissues, but the differences in gene expression levels were more significant, including 21 genes with higher expression levels and 27 genes with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

lower expression levels. Among them, the same subclasses of proteins did not have similar expression profiles, such as *AGO10* (*AGO10c* showed high expression in most tissues, while expression was significantly downregulated in mature leaf-jun08 and dehiscing hull, and *AGO10a* and *AGO10b* were mostly low and upregulated in some tissues) and *RDR1* (RDR1a showed high expression in most tissues, while expression was significantly downregulated in mature embryo, and the other eight *RDR1* genes, i.e., *RDR1b-i*, were mostly expressed at low levels) (Supplementary Table 7A and Figure 9). In addition, their potential stress responses (including the N response and heat stress response) were examined. At various concentrations of N supply, the expression of 16 RNAi-related genes in the root fraction was highest, while the expression of 14 RNAi-related genes remained low, and the expression levels of the remaining 18 genes were relatively intermediate. Notably, the expression levels of JrDRB7.1, *JrDCL2b*, and *JrRDR1i* increased with the change in N supply concentration; in particular, the expression levels of *JrRDR2* and *JrRDR1h* seemed to be significantly correlated with an increase in N concentration. In addition, RNAi-related genes were mostly heat-treated, resulting in restricted expression. Among them, approximately 60% of RNAi-related genes were highly expressed at 0 h of heat treatment, while they were approximately 30% higher after 2 h and 6 h of heat treatment. Interestingly, the expression of *AGO2* (*JrAGO2a* and *JrAGO2b*) proteins was induced by heat treatment (Supplementary Figure 4).

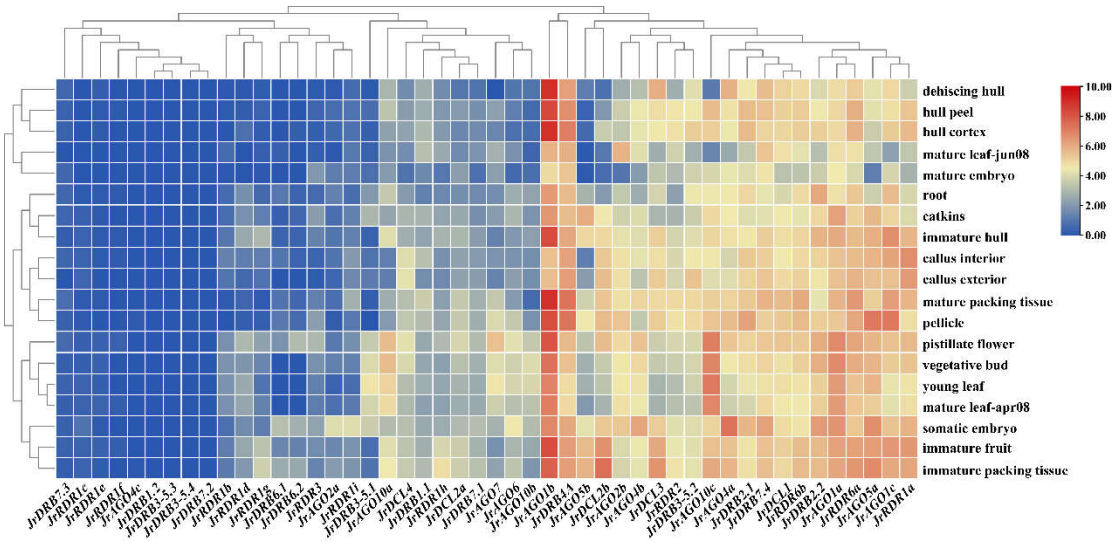


**Figure 9.** Heatmap of tissue-specific expression and abiotic stress expression of walnut RNAi-related genes. The colour bar represents the log2 expression level of each gene (TPM, transcripts per million). Colour bar annotation is included at the top of the image. The heatmap is coloured according to expression values, with blue, yellow and red representing low, medium and high transcription abundance, respectively.

### *3.9. Analysis of disease resistance expression of RNAi-related genes in walnut*

Due to the previous importance of RNAi-related genes for disease, it was analysed transcriptome data based on walnut induced by Colletotrichum gloeosporioides. The results showed that most of the walnut RNAi-related genes showed significant upregulation of expression compared to uninfested walnut leaves (fold change>2). Among them, more than 60% of genes were significantly upregulated under different infestation times, such as *JrAGO5b* and JrDRB*6.1*. Moreover, there were significant disease duration differences in the response of these RNAi-related genes to Colletotrichum gloeosporioides, with most genes showing upregulated expression levels at 24 h and 120 h and weaker expression upregulation levels at 48 h and 72 h. In particular, many homologous protein classes tended to be induced at the same level, such as *RDR1* and *RDR6* (Figure 10). Considering the complex network of RNAi-related genes with respect to the conditions for various tissues with stress expression, it is hypothesized that a large number of extended gene

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

families in walnut may be associated with stress and selected by functionalization (e.g., multifunctionalization and subfunctionalization) effects.

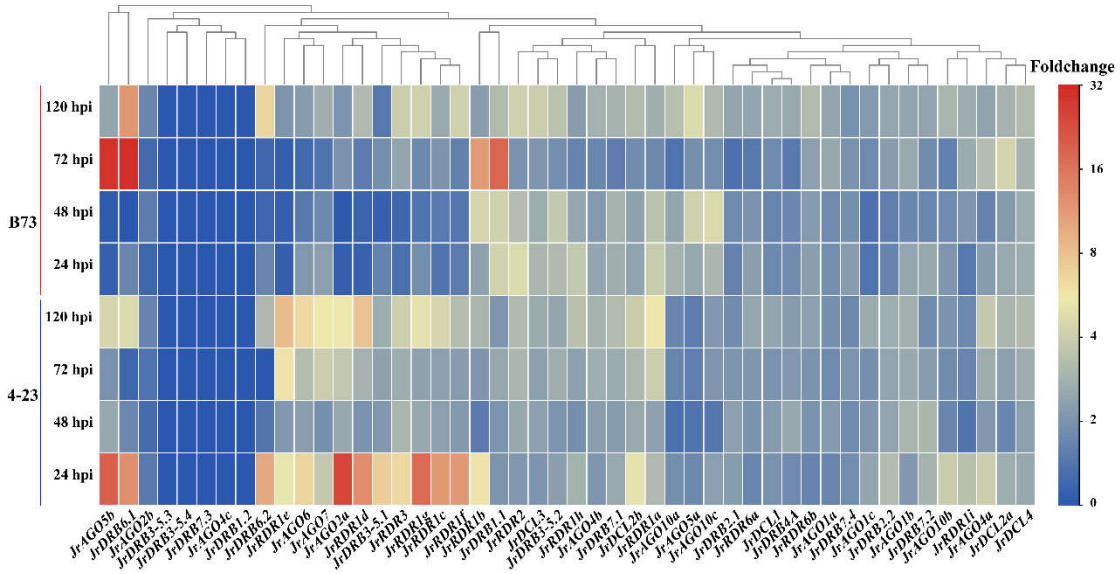


**Figure 10. Anti-Colletotrichum gloeosporioides expression analysis of walnut *AGO*, *RDR*, *DRB* and *DCL*.** The colour bar represents the log2 expression level of each gene (TPM, transcripts per million). Colour bar annotation is included at the top of the image. The heatmap is coloured according to expression values, with blue, yellow and red representing low, medium and high transcription abundance, respectively.

## 4. Discussion

Juglandaceae is an important source of timber and possesses many nutrient-rich dried fruit plants that play an important role in the development of the world agroforestry economy. However, in the walnut cultivation industry, many adversities, such as cold damage, pests and diseases, have seriously affected its quality and quantity production. In recent years, sequencing of many genomes of Juglandaceae plants has been completed (Guo et al., 2020). Evolutionary studies of their RNAi systems based on feasible methods are an important component to facilitate their production applications and help to preliminarily resolve natural disease resistance and gene functional differentiation in walnuts. For this purpose, the *DCL*, *RDR*, *DRB*, and *AGO* genes present in the walnut genome were comprehensively analysed and studied their structure, evolution, and possible interactions. The data were used to analyse the gene expression differentiation of RNAi-related genes, and these findings are expected to lay the foundation for walnut breeding research.

### *4.1. Evolution of RNAi-related genes in plants*

*DCL* genes can generally be classified as *DCL1*, *DCL2*, *DCL3*, and *DCL4*, and the copy number is generally no more than three (Chow et al., 2019; Sabbione et al., 2019; Boland et al., 2011; Parvez Mosharaf et al., 2019; Kakiyama et al., 2019; Katsarou et al., 2016; Dubey et al., 2020; Moura et al., 2019). The walnut *DCL* gene also includes the above features, and only *DCL2* produces two copies of the walnut *DCL* gene. In fact, *DCL2* and *DCL3* may be copied in the *DCL* family, including multicopy *DCL2* (*DCL2a* and *DCL2b*) reported from poplar, rice and maize, and cotton plants (Moura et al., 2019). Meanwhile, we also have this multicopy feature in the walnut family, even up to four copies in Juglans nigra. In addition, *DCL2* duplication may be associated with independent evolution of monocotyledons and true dicotyledons after differentiation (Margis et al., 2006); in the collinearity feature section of the text, the collinearity of the *JrDCL2b* gene was found only in nonherbaceous plants in dicotyledons, suggesting that the evolutionary process of this gene may depend on the evolution of plant stem organs. Second, previous results suggest that *DCL3* replication may be a unique event in eudicots (Moura et al., 2019), but we suggest that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

*DCL3* replication may be heavily dependent on the formation of some species-specific branches rather than a universal replication event in eudicots, since the Juglandaceae explored here are all prepared with only a single *DCL3*.

A total of six *RDR* genes have been identified in Arabidopsis, but in more studies, it has been shown that *RDR4* and *RDR5* may be copies of *RDR3*, thus delineating the four categories of *RDR1*, *RDR2*, *RDR3*, and *RDR6*. Walnut *RDR* genes contain four classes of *RDR1*, *RDR2*, *RDR3*, and *RDR6* with a total of 13, of which *RDR6* copy number is 2, while *RDR2* and *RDR3* are single copies, which is similar to the structure in most studies. However, the *RDR1* copy number of 9 is unique, seven of which are arranged in clusters on chromosome 16, which is much more than the one to three in the rest of the species (Cui et al., 2020; Sabbione et al., 2019; Boland et al., 2011; Prakash & Chakraborty, 2019; Yan et al., 2019a; Potter et al., 2018). Interestingly, aberrant expansions of *RDR* have long been noted in polyploid species and wild species, but it appears that there are clear differences in the evolutionary patterns of *RDR* in different taxa of species (Cao et al., 2016). For example, 3 and 5 copies of *RDR3*/*4*/*5* are present in Arabidopsis and European oilseed rape, while the massive expansion of *RDR1* in the walnut genome does not seem to be reflected in others, and the massive expansion of *RDR1* is present in a variety of Juglandaceae species.

The *AGO* gene family is a typical multicopy family classified into three classes and 11 subfamilies (I: *AGO1*, *AGO5*, *AGO10*; II: *AGO2*, *AGO3*, *AGO7*; III: *OsAGO15*, *AGO9*, *AGO8*, *AGO6*, *AGO4*). Walnut *AGO* genes were distributed in seven subclasses of the three classes (I: *AGO1*, *AGO5*, *AGO10*; II: *AGO2*/*3*, *AGO7*; III: *AGO6*, *AGO4*) with copy numbers of 1 to 3. These results were similar to most species (El-Gebali et al., 2019; Wang et al., 2017; Curaba & Chen, 2008; Sabbione et al., 2019; Das et al., 2020; Kobayashi & Tomari, 2020; Ciechanowska et al., 2021; Wang et al., 2012). Interestingly, while *DCL* and *RDR* have produced a large number of expansions unique to Juglandaceae, there has not been a large expansion of genes in *AGO*, where the number of walnuts is 15 and the number of other Juglandaceae plants is mostly also in the range of 12-17.

*DRB* mainly consists of four subclasses (*DRB1*, *DRB2-3-5*, *DRB4*, *DRB6*), where *DRB2* and *DRB7* are sometimes considered independent branches (Clavel et al., 2016). However, no matter how to slice the *DRB*, the *DRB* gene of walnut is twofold expanded among various subbranches compared to Arabidopsis. Meanwhile, substantial expansion of *DRB1* and *DRB3* remained in other walnut plants, but other branches were lost and selectively expanded.

### *4.2. Structure of RNAi-related genes in plants*

Walnut has seven core and two variable structural domains, of which PAZ confers DCL binding ability to RNA, and RNase III (Ribonuclease_3), Helicase_C, DEAD is the endonuclease active domain of DCL enzymes, , which enhance DCL's ability to deconvolve double-stranded molecules (Ciechanowska et al., 2021; Macrae et al., 2006). In contrast, the dsrm and DND1_DSRM structural domains are only distributed within the JrDCL1 and JrDCL4 proteins, which may lead to enhanced RNA binding capacity of DCL1 and DCL4, resulting in functional differences with DCL2s and DCL3. Walnut RDR enzymes exist only in an RdRP structural domain of variable length, which is the core region where RDR exercises its function (Curaba & Chen, 2008; Wang et al., 2020).

Walnut AGO proteins possess seven structural domains, five of which are shared, and the remaining ArgoMid is distributed in the class I proteins of walnut AGO. Gly-rich_Ago1 is only distributed in walnut AGO1s protein, and PAZ and PIWI are key to the formation of the core active region of AGO, and the simultaneous presence of ArgoN, ArgoMid, and PIWI would be more helpful for the formation of AGO functional regions. Thus, these specifically distributed structures may lead to functional differences in different nucleolar AGO proteins (Boland et al., 2011; Carbonell & Carrington, 2015; Zhang et al., 2014; Muller et al., 2019). DRB proteins include one or more DSRM (double-stranded RNA binding motif) structures, a result related to their hydrolytic activity on dsRNA

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

(Montavon et al., 2017). Walnut DRB enzymes are present in one or more DSRM structural domains, which are the key regions that guarantee their hydrolytic activity on dsRNA (Curaba & Chen, 2008; Wang et al., 2020). Among them, walnuts possess 27%, 66%, and 5% of the proteins possessing one, two, or three DSRM structures, respectively. Meanwhile, other walnut family plants possessing 1-4 DSRM structure proteins are abundantly present, but the proteins with 2 DSRM structures remain the most abundant. Second, the Methyltransf_11 structure was also found in the JrDRB7.1 protein, a result that may have a unique role.

Different RDR-DCL/DRB-AGO combinations are synergistically involved in specialized RNA silencing to control invading nucleic acids from endogenous (mainly transposon) or exogenous (mainly viral) sources and are mediated by a variety of sRNAs, such as miRNA, trans siRNA (ta-siRNA), natural antisense transcript-siRNA (nat-siRNA), and viral-derived siRNA (vsiRNA) (Fang et al., 2020). In the present study, rich PPI network analysis revealed a large number of interactions among these four core components, which may mediate the multiple RNA silencing pathways in which they may be involved. Notably, the massive expansion of various RNAi-related genes in walnut also interacts positively with components of other families.

*4.3. Function of RNAi-related genes in plants*

The RNAi pathway effectively mediates plant adaptation to the outside world and relies on the small RNAs produced to further regulate plant resistance (Chow et al., 2019; Hung & Slotkin, 2021). Available evidence suggests that the RNAi pathway is mainly accomplished by three classes of proteins, DCL, RDR, and AGO, with DCL2, DCL4, RDR2, RDR6, AGO1, AGO2, AGO3, AGO4, AGO5, AGO7, AGO10, and AGO18 being the most important for plant RNAi resistance effects, while other DCL1 and DCL3 led by proteins appear to exist to mediate small RNA formation as well as enrich neofunctionalization (Fang et al., 2020; Parent et al., 2015; Katsarou et al., 2016; Kakiyama et al., 2019). Second, DRB proteins can also effectively complement many aspects of the system (Clavel et al., 2016).

Various gene families may often be somewhat selective for the type of tissue, as is the case for RNAi-associated gene families (Hsu et al., 2021; Cao et al., 2016; Sabbione et al., 2019; Yan et al., 2019b). At the same time, these genes that produce tissue specificity can often also produce important effects on specific tissue life activities or tissue resistance. Walnut RNAi-associated genes show unique tissue specificity for different tissues. Moreover, the same protein subclasses do not have similar expression profiles, such as *AGO10* and *RDR1*. This finding greatly reflects the unique functional differentiation of RNAi and its potential resistance to pathogens (infected tissues may be richer in resistance genes).

Indeed, many different gene families in many species have produced rich functional differentiation, leading to a large functional expansion of enriched genes, e.g., MADS and cytochrome P450 (Schilling et al., 2020; Fang et al., 2021b). Due to the unique tissue expression of RNAi genes and the low tissue expression of a small number of genes, we speculate that stress may induce differential expression of these RNAi-related genes. In fact, heat stress and nitrogen stress are two completely different types of stresses (environmental factor forcing and nutritional stress). The results showed that the expression of 16 RNAi-related genes was highest in the root fraction, and the expression of *JrDRB7.1*, *JrDCL2b*, and *JrRDR1i* increased with N supply concentration. Most of the RNAi-related genes were heat-treated, resulting in restricted expression, although a small number of proteins, including *AGO2* (*JrAGO2a* and *JrAGO2b*), could be upregulated by heat treatment-induced upregulated expression. This suggests that the RNAi-associated gene family produced significant functional differentiation, and in particular, some taxa of proteins may develop unique response characteristics to different tissues or stresses.

RNA silencing protects plants from viral infection (Ding & Voinnet, 2007). This antiviral immunity involves the production of virus-derived small interfering RNA (viRNA) and leads to virus-specific silencing through viRNA-induced effector complexes. In recent

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

studies, plant miRNAs were shown to be differentially expressed upon inoculation with fungal pathogens, such as *Bradyrhizobium* (Xin et al., 2010), *Fusarium virguliforme* (Radwan et al., 2011), *Verticillium dahliae* (Yang et al., 2013), *Verticillium longisporum* (Shen et al., 2014), *Magnaporthe oryzae* (Li et al., 2014) and *Botrytis cinerea* (Jin & Wu, 2015). More importantly, mutants of key components of the RNA silencing machinery exhibit altered susceptibility to fungal pathogens, including two Verticillium species (Shen et al., 2014; Ellendorff et al., 2009). In addition, *Botrytis cinerea* has been reported to suppress plant defences by hijacking the host RNA interference pathway to produce small RNAs (Weiberg et al., 2013). Therefore, it is crucial to clarify the major contribution of these components to biotic stress, and transcriptome data from walnut induced by *C. gloeosporioides* showed that the expression of most walnut RNAi-related genes was significantly upregulated (fold change>2). In addition, many homologous genes tended to be induced at the same level, such as *RDR1* and *RDR6*. In conclusion, this result indicates that RNAi-related genes still have an important role in the expression of disease resistance traits in walnut and further exploration of their specific resistance roles. In conclusion, this result indicates that RNAi-related genes still have an important role in the expression of disease resistance characteristics in walnut, and it is important to further explore their specific roles and resistance breeding.

## 5. Conclusion

In summary, walnut *AGO*, *DCL*, *DRB* and *RDR* gene families were identified for the first time, completed an analysis of their sequence characteristics, evolutionary features (within species, within genus, within family, between woody species, within dicotyledons, between different orders at multiple levels, respectively) and interaction possibilities, and finally, characterized the *AGO*, *DCL*, *DRB* and *RDR* genes based on transcriptome. Finally, based on the transcriptome, the potential functional was also characterized from differentiation of *AGO*, *DCL*, *DRB* and *RDR* genes These findings are expected to lay the foundation for the study of breeding research in walnut and have reference value for the evolutionary study of plant resistance.

**Supplementary Materials:** Supplementary Figure 1 Distribution of the structural domains of other Juglandaceae RDR (A), DRB (B), DCL (C) and AGO (D); Supplementary Figure 2 Evolutionary characteristics of AGO, RDR, DRB and DCL homologous proteins of other Juglandaceae plants with walnut. (A) Analysis of covariance between walnut and two Juglandaceae plants; analysis of selection pressure and differentiation time of (B) collinearity genes and (C) similar proteins between walnut and nine Juglandaceae plants; Supplementary Figure 3 Promoter analysis of the walnut *AGO*, *RDR*, *DRB* and *DCL* genes. Rootless trees were generated by the neighbour-joining (NJ) method using MEGA7 software. Numbers next to branches indicate 1,000 bootstrap replicates as a percentage. Different promoters are indicated by rectangles of different colours to represent homeopathic elements. Detailed comments are included in the top right panel; Supplementary Figure 4 Walnut *AGO*, *RDR*, *DRB* and *DCL* nitrogen (A) and heat (B) stress treatment expression analysis; Supplementary Table 1 RNAi-related protein sequences of other species used; Supplementary Table 2 Collinearity of *AGO*, *RDR*, *DRB* and *DCL* between different plants and walnuts; Supplementary Table 3 Basic information of other Juglandaceae AGO, RDR, DRB and DCL; Supplementary Table 4 Selection pressure and differentiation time of congeners of Juglandaceae *AGO*, *RDR*, *DRB* and *DCL*; Supplementary Table 5 Promoter analysis of *AGO*, *RDR*, *DRB* and *DCL* between different plants and walnut; Supplementary Table 6 Basic information on the *AGO*, *RDR* and *DCL* genes in walnut. Subcellular localization represents only the outcome with the highest confidence of the predicted outcome; Supplementary Table 7 Walnut *AGO*, *RDR*, *DRB* and *DCL* tissue and stress expression analysis.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China (32060673), Guizhou Provincial Science and Technology Projects (Qian Ke He Ji Chu-ZK [2022] General 071) and projects of Guizhou University (No. Gui Da Pei Yu[2019]52) co-funded.

**Conflicts of Interest:** The authors declare that they have no competing interests.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1

## References


1. Bai M, Yang GS, Chen WT, Mao ZC, Kang HX, Chen GH, Yang YH, Xie BY, 2012. Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene* **501**, 52-62.
2. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS, 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**, W202-W208.
3. Baksa I, Szittya G, 2017. Identification of ARGONAUTE/Small RNA Cleavage Sites by Degradome Sequencing. *Methods Mol Biol* **1640**, 113-128.
4. Boland A, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O, 2011. Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc Natl Acad Sci U S A* **108**, 10466-10471.
5. Cao JY, Xu YP, Li W, Li SS, Rahman H, Cai XZ, 2016. Genome-Wide Identification of Dicer-Like, Argonaute, and RNA-Dependent RNA Polymerase Gene Families in *Brassica* Species and Functional Analyses of Their Arabidopsis Homologs in Resistance to *Sclerotinia sclerotiorum*. *Front Plant Sci* **7**, 1614.
6. Carbonell A, Carrington JC, 2015. Antiviral roles of plant ARGONAUTES. *Curr Opin Plant Biol* **27**, 111-117.
7. Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R, 2020. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant* **13**, 1194-1202.
8. Chow FW, Koutsovoulos G, Ovando-Vazquez C, Neophytou K, Bermúdez-Barriento JR, Laetsch DR, Robertson E, Kumar S, Claycomb JM, Blaxter M, Abreu-Goodger C, Buck AH, 2019. Secretion of an Argonaute protein by a parasitic nematode and the evolution of its siRNA guides. *Nucleic Acids Research* **47**, 3594-3606.
9. Ciechanowska K, Pokornowska M, Kurzynska-Kokorniak A, 2021. Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases. *International Journal of Molecular Sciences* **22**.
10. Clavel M, Pelissier T, Montavon T, Tschopp M, Pouch-Pélissier M, Descombin J, Jean V, Dunoyer P, Bousquet-Antonelli C, Deragon J, 2016. Evolutionary history of double-stranded RNA binding proteins in plants: identification of new cofactors involved in easiRNA biogenesis. *Plant Molecular Biology* **91**, 131-147.
11. Cui DL, Meng JY, Ren XY, Yue JJ, Fu HY, Huang MT, Zhang QQ, Gao SJ, 2020. Genome-wide identification and characterization of *DCL*, *AGO* and *RDR* gene families in *Saccharum spontaneum*. *Sci Rep* **10**, 13202.
12. Curaba J, Chen X, 2008. Biochemical activities of Arabidopsis RNA-dependent RNA polymerase 6. *J Biol Chem* **283**, 3059-3066.
13. Das S, Swetha C, Pachamuthu K, Nair A, Shivaprasad PV, 2020. Loss of function of *Oryza sativa* Argonaute 18 induces male sterility and reduction in phased small RNAs. *Plant Reprod* **33**, 59-73.
14. Ding SW, Voinnet O, 2007. Antiviral immunity directed by small RNAs. *Cell* **130**, 413-26.
15. Dubey H, Kiran K, Jaswal R, *et al*., 2020. Identification and characterization of Dicer-like genes in leaf rust pathogen (*Puccinia triticina*) of wheat. *Funct Integr Genomics* **20**, 711-721.
16. Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C, 2021. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Research* **49**, W216-W227.
17. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD, 2019. The Pfam protein families database in 2019. *Nucleic Acids Research* **47**, D427-D432.
18. Ellendorff U, Fradin EF, De Jonge R, Thomma BPHJ, 2009. RNA silencing is required for Arabidopsis defence against Verticillium wilt disease. *Journal of Experimental Botany* **60**, 591-602.
19. Esposito S, Aversano R, D'amelia V, Villano C, Alioto D, Mirouze M, Carputo D, 2018. Dicer-like and RNA-dependent RNA polymerase gene family identification and annotation in the cultivated *Solanum tuberosum* and its wild relative *S. commersonii*. *Planta* **248**, 729-743.
20. Fang H, Liu X, Dong Y, Feng S, Zhou R, Wang CX, Ma XM, Liu JN, Yang KQ, 2021a. Transcriptome and proteome analysis of walnut (*Juglans regia* L.) fruit in response to infection by *Colletotrichum gloeosporioides*. *BMC Plant Biol* **21**, 249.
21. Fang Y, Jiang J, Du Q, Luo L, Li X, Xie X, 2021b. Cytochrome P450 Superfamily: Evolutionary and Functional Divergence in Sorghum (*Sorghum bicolor*) Stress Resistance. *J Agric Food Chem* **69**, 10952-10961.
22. Fang Y, Li Y, Yue N, Zhao ZB, Yang ZF, Wang Yong, Long YH, 2020. Advance in the antiviral mechanism of RNA interference. *Journal of Mountain Agriculture and Biology* **39**, 50-56.
23. Feng S, Fang H, Liu X, Dong Y, Wang Q, Yang KQ, 2021. Genome-wide identification and characterization of long non-coding RNAs conferring resistance to *Colletotrichum gloeosporioides* in walnut (*Juglans regia*). *BMC Genomics* **22**, 15.
24. Gan D, Zhan M, Yang F, Zhang QQ, Hu KL, Xu WJ, Lu QH, Zhang L, Liang DD, 2017. Expression analysis of argonaute, Dicer-like, and RNA-dependent RNA polymerase genes in cucumber (*Cucumis sativus* L.) in response to abiotic stress. *J Genet* **96**, 235-249.
25. Gu Z, Cavalcanti A, Chen F-C, Bouman P, Li W-H, 2002. Extent of gene duplication in the genomes of Drosophila, nematode, and yeast. *Molecular Biology and Evolution* **19**, 256–262.
26. Guo W, Chen J, Li J, Huang J, Wang Z, Lim KJ, 2020. Portal of Juglandaceae: A comprehensive platform for Juglandaceae study. *Hortic Res* **7**, 35.
27. Hartman E, Wang Z, Zhang Q, Roy K, Chanfreau G, Feigon J, 2013. Intrinsic dynamics of an extended hydrophobic core in the S. cerevisiae RNase III dsRBD contributes to recognition of specific RNA binding sites. *J Mol Biol* **425**, 546-562.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1


28. Hsu HF, Chen WH, Shen YH, Hsu WH, Mao WT, Yang CH, 2021. Multifunctional evolution of B and *AGL6* MADS box genes in orchids. *Nature Communications* **12**, 902.
29. Huang Z, Duan W, Song X, Tang J, Wu P, Zhang B, Hou XL, 2015. Retention, Molecular Evolution, and Expression Divergence of the Auxin/Indole Acetic Acid and Auxin Response Factor Gene Families in Brassica Rapa Shed Light on Their Evolution Patterns in Plants. *Genome Biol Evol* **8**, 302-16.
30. Hung YH, Slotkin RK, 2021. The initiation of RNA interference (RNAi) in plants. *Curr Opin Plant Biol* **61**, 102014.
31. Jin WB, Wu FL, 2015. Characterization of miRNAs associated with *Botrytis cinerea* infection of tomato leaves. *Bmc Plant Biology* **15**.
32. Kakiyama S, Tabara M, Nishibori Y, Moriyama H, Fukuhara T, 2019. Long DCL4-substrate dsRNAs efficiently induce RNA interference in plant cells. *Sci Rep* **9**, 6920.
33. Katsarou K, Mavrothalassiti E, Dermauw W, Van Leeuwen T, Kalantidis K, 2016. Combined Activity of DCL2 and DCL3 Is Crucial in the Defense against *Potato Spindle Tuber Viroid*. *PLoS Pathog* **12**, e1005936.
34. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL, 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915.
35. Kobayashi H, Tomari Y, 2020. Identification of an AGO (Argonaute) protein as a prey of TER94/VCP. *Autophagy* **16**, 190-192.
36. Kumar S, Stecher G, Tamura K, 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **33**, 1870-1874.
37. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, de Peer YV , Rouzé P, Rombauts S, 2002. Plant CARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* **30**, 325-327.
38. Letunic I, Bork P, 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research* **46**, D493-D496.
39. Li Y, Lu YG, Shi Y, Wu L, Xu YJ, Huang F, Guo XY, Zhang Y, Fan J, Zhao JQ, Zhang HY, Xu PZ, Zhou JM, Wu XJ, Wang PR, Wang WN, 2014. Multiple Rice MicroRNAs Are Involved in Immunity against the Blast Fungus *Magnaporthe oryzae*. *Plant Physiology* **164**, 1077-1092.
40. Li ZC, Li WQ, Guo MX, Liu SM, Liu L, Yu Y, Mo BX, Chen XM, Gao L, 2022. Origin, evolution and diversification of plant ARGONAUTE proteins. *Plant Journal* **109**, 1086-1097.
41. Liao Y, Smyth GK, Shi W, 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930.
42. Macrae IJ, Li F, Zhou K, Cande WZ, Doudna JA, 2006. Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb Symp Quant Biol* **71**, 73-80.
43. Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, E. Finnegan J, Waterhouse PM, 2006. The evolution and diversification of Dicers in plants. *FEBS Lett* **580**, 2442-50.
44. Mcwhite CD, Papoulas O, Drew K, Cox RM, June V, Dong OX, Kwon T, Wan CH, Salmi ML, Roux SJ , Browning KS, Chen ZJ, Ronald PC, Marcotte EM, 2020. A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies. *Cell* **181**, 460-74 e14.
45. Mochizuki K, Gorovsky MA, 2005. A Dicer-like protein in Tetrahymena has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev* **19**, 77-89.
46. Montavon T, Kwon Y, Zimmermann A, Michel F, Dunoyer P, 2017. New DRB complexes for new DRB functions in plants. *RNA Biology* **14**, 1637-1641.
47. Moura MO, Fausto AKS, Fanelli A, Guedes FA de F, Silva T da F, Romanel E, Vaslin MFS, 2019. Genome-wide identification of the Dicer-like family in cotton and analysis of the *DCL* expression modulation in response to biotic stress in two contrasting commercial cultivars. *Bmc Plant Biology* **19**.
48. Mukherjee K, Campos H, Kolaczkowski B, 2013. Evolution of Animal and Plant Dicers: Early Parallel Duplications and Recurrent Adaptation of Antiviral RNA Binding in Plants. *Molecular Biology and Evolution* **30**, 627-641.
49. Muller M, Fazi F, Ciaudo C, 2019. Argonaute Proteins: From Structure to Function in Development and Pathological Cell Fate Determination. *Front Cell Dev Biol* **7**, 360.
50. Parent JS, Bouteiller N, Elmayan T, Vaucheret H, 2015. Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing. *Plant J* **81**, 223-232.
51. Parvez Mosharaf M, Akond Z, Hadiul Kabir M, Nurul Haque Mollah M, 2019. Genome-wide identification, characterization and phylogenetic analysis of Dicer-like (DCL) gene family in *Coffea arabica*. *Bioinformation* **15**, 824-31.
52. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD, 2018. HMMER web server: 2018 update. *Nucleic Acids Research* **46**, 200-204.
53. Prakash V, Chakraborty S, 2019. Identification of transcription factor binding sites on promoter of RNA dependent RNA polymerases (RDRs) and interacting partners of RDR proteins through in silico analysis. *Physiol Mol Biol Plants* **25**, 1055-1071.
54. Qin L, Mo N, Muhammad T, Liang Y, 2018. Genome-Wide Analysis of *DCL*, *AGO*, and *RDR* Gene Families in Pepper (*Capsicum Annuum* L.). *International Journal of Molecular Sciences* **19**.
55. Radwan O, Liu Y, Clough SJ, 2011. Transcriptional Analysis of Soybean Root Response to *Fusarium virguliforme*, the Causal Agent of Sudden Death Syndrome. *Molecular Plant-Microbe Interactions* **24**, 958-972.
56. Sabbione A, Daurelio L, Vegetti A, Talon M, Tadeo F, Dotto M, 2019. Genome-wide analysis of *AGO*, *DCL* and *RDR* gene families reveals RNA-directed DNA methylation is involved in fruit abscission in *Citrus sinensis*. *BMC Plant Biol* **19**, 401.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2022 doi:10.20944/preprints202210.0323.v1


57. Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R, 2020. Genome-wide analysis of *MIKC-type MADS-box* genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist* **225**, 511-529.
58. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T, 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504.
59. Shen D, Suhrkamp I, Wang Y, Liu SY, Menkhaus J, Verreet JA, Fan LJ, Cai DG, 2014. Identification and characterization of microRNAs in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using *Brassica* AA (*Brassica rapa*) and CC (*Brassica oleracea*) as reference genomes. *New Phytologist* **204**, 577-594.
60. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH, 2019. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Research* **47**, 270-275.
61. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C, 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research* **45**, 362-368.
62. Tang K, Duan CG, Zhang H, Zhu JK, 2017. Computational Analysis of Genome-Wide ARGONAUTE-Dependent DNA Methylation in Plants. *Methods Mol Biol* **1640**, 219-225.
63. Wang B, Zhang J, Pei D, Yu L, 2021. Combined effects of water stress and salinity on growth, physiological, and biochemical traits in two walnut genotypes. *Physiol Plant* **172**, 176-187.
64. Wang M, Deng Y, Shao F, *et al*., 2017. ARGONAUTE Genes in Salvia miltiorrhiza: Identification, Characterization, and Genetic Transformation. *Methods Mol Biol* **1640**, 173-189.
65. Wang M, Li R, Shu B, Jing X, Ye HQ, Gong P, 2020. Stringent control of the RNA-dependent RNA polymerase translocation revealed by multiple intermediate structures. *Nature Communications* **11**, 2605.
66. Wang YP, Tang HB, Debarry JD, Tan X, Li JP, Wang XY, Lee TH, Jin HZ, Marler B, Guo H, Kissinger JC, Paterson AH, 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40**.
67. Weiberg A, Wang M, Lin FM, Zhao HW, Zhang ZH, Kaloshian I, Huang HD, Jin HL, 2013. Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science* **342**, 118-123.
68. Xin MM, Wang Y, Yao YY, Xie CJ, Peng HR, Ni ZF, Sun QX, 2010. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). *Bmc Plant Biology* **10**.
69. Yan F, Li H, Zhao P, 2019a. Genome-Wide Identification and Transcriptional Expression of the *PAL* Gene Family in Common Walnut (*Juglans Regia* L.). *Genes (Basel)* **10**.
70. Yan F, Zhou HJ, Yue M, Yang G, Li HZ, Zhang SX, Zhao P, 2019b. Genome-Wide Identification and Transcriptional Expression Profiles of the *F-box* Gene Family in Common Walnut (*Juglans regia* L.). *Forests* **10**.
71. Yang L, Jue DW, Li W, Zhang RJ, Chen M, Yang Q, 2013. Identification of miRNA from Eggplant (*Solanum melongena* L.) by Small RNA Deep Sequencing and Their Response to *Verticillium dahliae* Infection. *PLoS One* **8**.
72. Zhang H, Xia R, Meyers BC, Walbot V, 2015. Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Curr Opin Plant Biol* **27**, 84-90.
73. Zhang J, Zhang W, Ji F, Qiu J, Song XB, Bu DC, Pan G, Ma QG, Chen JX, Huang RM, Chang YY, Pei D, 2020. A high-quality walnut genome assembly reveals extensive gene expression divergences after whole-genome duplication. *Plant Biotechnol J* **18**, 1848-1850.
74. Zhang X, Niu D, Carbonell A, Wang A, Lee A, Tun V, Wang ZH, Carrington JC, Chang CA, Jin HL, 2014. ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nature Communications* **5**, 5468.
75. Zhao HL, Zhao K, Wang J, Chen X, Chen Z, Cai RH, Xiang Y, 2015a. Comprehensive Analysis of Dicer-Like, Argonaute, and RNA-dependent RNA Polymerase Gene Families in Grapevine (*Vitis Vinifera*). *Journal of Plant Growth Regulation* **34**, 108-21.
76. Zhao K, Zhao H, Chen Z, Feng L, Ren J, Cai RH, Xiang Y, 2015b. The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution. *J Genet* **94**, 317-321.
77. Zong J, Yao X, Yin JY, Zhang DB, Ma H, 2009. Evolution of the RNA-dependent RNA polymerase (*RdRP*) genes: Duplications and possible losses before and after the divergence of major eukaryotic groups. *Gene* **447**, 29-39.