Article

Genome-Wide Identification and Coexpression Network Analysis of DNA Methylation Pathway Genes and Their Differentiated Functions in *Ginkgo biloba* L.

Caiyun Gao , Miao Deng, Xiaoming Yang , Wanwen Yu, Jinfeng Cai, Yuanbao Shi, Zhibo Zhu, Tingting Zhou, Liangjiao Xue, Fuliang Cao , Guibin Wang and Fang-Fang Fu *

Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing 210037, China; gaocaiyunnanlin@njfu.edu.cn (C.G.); mdeng@njfu.edu.cn (M.D.); xmyang@njfu.edu.cn (X.Y.); youeryuww@163.com (W.Y.); caijinfeng1984@126.com (J.C.); ybao_shi@163.com (Y.S.); vuezhu@gmail.com (Z.Z.); tingtingzhou@njfu.edu.cn (T.Z.); lxue@njfu.edu.cn (L.X.); fuliangcaonjfu@163.com (F.C.); gbwang@njfu.com.cn (G.W.)

* Correspondence: fffu@njfu.edu.cn; Tel.: +86-8542-7921

Received: 17 August 2020; Accepted: 5 October 2020; Published: 9 October 2020

Abstract: DNA methylation plays a vital role in diverse biological processes. DNA methyltransferases (DNMTs) genes and RNA-directed DNA methylation (RdDM)-related genes are key genes responsible for establishing and maintaining genome DNA methylation in plants. In the present study, we systematically identified nine *GbDNMTs* in *Ginkgo biloba*, including the three common families of *GbMET1a/1b*, *GbCMT2*, and *GbDRMa/b/2a/2b/2c*, and a fourth family—*GbDNMT3*—which is absent in most angiosperms. We also identified twenty RdDM-related genes, including four *GbDCLs*, six *GbAGOs*, and ten *GbRDRs*. Expression analysis of the genes showed the different patterns of individual genes, and 15 of 29 genes displayed expression change under five types of abiotic stress. Gene coexpression analysis and weighted gene co-expression network analysis (WGCNA) using 126 public transcriptomic datasets revealed that these genes were clustered into two groups. In group I, genes covered members from all six families which were preferentially expressed in the ovulate strobile and fruit. A gene ontology (GO) enrichment analysis of WGCNA modules indicated that group I genes were most correlated with the biological process of cell proliferation. Group II only consisted of RdDM-related genes, including *GbDRMs*, *GbAGOs*, and *GbRDRs*, but no *GbDCLs*, and these genes were specifically expressed in the cambium, suggesting that they may function in a dicer-like (DCL)-independent RdDM pathway in specific tissues. The gene module related to group II was most enriched in signal transduction, cell communication, and the response to the stimulus. These results demonstrate that gene family members could be conserved or diverged across species, and multi-member families in the same pathway may cluster into different modules to function differentially. The study provides insight into the DNA methylation genes and their possible functions in *G. biloba*, laying a foundation for the further study of DNA methylation in gymnosperms.

Keywords: DNA methylation; gene family; WGCNA; *Ginkgo*

1. Introduction

Cytosine DNA methylation is an epigenetic modification that regulates gene expression [1,2]. In the plant kingdom, 5mC is usually methylated in three cytosine contexts: CG, CHG, and CHH (H = A/T/C). These are mainly established by RNA-directed DNA methylation (RdDM) and maintained by different DNA methyltransferase (DNMT) pathways, including methyltransferase 1 (MET1), chromomethylase

(CMT), domains rearranged methyltransferase (DRM), and DNA methyltransferase 3 (DNMT3) [1–5]. MET1s are homologs to mammalian DNMT1 and maintain CG methylation. The plant-specific CMTs play an important role in maintaining CHH, CHG, and the heterochromatin status. The DRM enzymes (DRM2s) that combine with the RdDM pathway are mainly responsible for de novo methylation at CHH sites [5,6]. DNMT3s are ancient DNA methyltransferases (DNMTs) that were overlooked in plants until being identified in *Physcomitrella patens* in the last year [7]. Dicer-likes (DCLs), argonautes (AGOs), and RNA-dependent RNA polymerases (RDRs) are the main components in the RdDM pathway [1]. RDRs are responsible for synthesizing double-stranded RNA (dsRNA) from the single-stranded RNA template. DCLs cleave the dsRNA into 21–24 nt long small RNAs, which are then loaded into an AGO protein-containing RNA-induced silencing complex (RISC) to direct gene silencing [8–10].

DNA methylation plays an important role, not only in growth and development regulation, but also in response to various stresses [11]. For example, low levels of DNA methylation lead to delayed flowering in *Arabidopsis* [12]. However, high levels of DNA methylation play a critical role in orange fruit development and ripening [13]. In *Citrus sinensis*, the important components of the RdDM pathway are down-regulated in the fruit abscission zone [8]. In apple, low DNA methylation levels are associated with smaller fruits [14]. Moreover, DNA methylation levels are different at different stages of tissues. In peach, the shoot apical meristem of young seedlings has a higher DNA methylation level than in adults [15]. In *Daucus carota*, the levels of DNA methylation in adult leaves are higher than in seedlings. However, there are higher DNA methylation levels in the roots of seedlings than adults [16]. The different DNA methylation levels in various tissues may be related to different DNA methylation pathways [17]. The above facts demonstrate that DNA methylation plays a pivotal role in growth and development. Moreover, DNA methylation-related genes are involved in various stresses. For example, most DNMTs were significantly down-regulated under cold and drought stresses in the tea plant [3]. Additionally, the expression levels of DNMTs changed greatly under abiotic stress in rapeseed [2]. Moreover, RdDM-related genes exhibited distinct expression patterns under different stress conditions in various plants [9,10,18–20].

Weighted gene co-expression network analysis (WGCNA) is a method that can identify gene modules in expression levels of highly correlated genes among a large number of samples [21]. Previous researchers utilized the public gene expression data of rice to construct a gene co-expression network and employed WGCNA to identify the modules in the gene co-expression network, applying it as a source of functional annotation for rice genes [22]. WGCNA, as a popular technique, has been widely used in *Arabidopsis*, maize, poplar, soybean, and other plants [22–25].

Ginkgo biloba L. is a gymnosperm species and considered as a “live fossil” in the plant kingdom, which has great economic, ecological, and medicinal values [26–29]. In recent years, the intensification of an extreme climate has had a great influence on the growth and development of plants. More and more research has focused on the expression levels of *G. biloba* genes at different growth stages or under various environmental stress conditions [27,30–35]. However, DNA methylation related research has scarcely involved *G. biloba*, despite the fact that its potential importance has been proven in wood species. In the present study, we first identified and characterized nine DNMTs and 20 RdDM-related genes in the *G. biloba* genome, respectively. Their conserved domains, motifs, gene structure, phylogenetic analysis, chromosomal location, and expression profiles in various tissues and under abiotic stresses were investigated. Furthermore, a large number of public RNA-seq data of *G. biloba* were used to explore the function of DNA methylation pathway genes more deeply. We obtained high-quality RNA-seq data of 126 samples from the NCBI Sequence Read Archive (SRA) database and assigned them to 18 modules with WGCNA. We further identified modules highly related to DNA methylation pathway genes, which were used for exploring their functions through gene ontology (GO) enrichment analysis [36]. Our study provides valuable information for further understanding the functional roles of DNA methylation in *G. biloba*.

2. Materials and Methods

2.1. Identification of DNA Methylation Pathway Gene Families in *G. biloba*

The genome sequences of *G. biloba* were downloaded from <http://gigadb.org/dataset/100613>. The hidden Markov model (HMM) profiles of DNA methylation pathway protein domains were extracted from the Pfam database (<http://pfam.xfam.org/>) and then used to search for corresponding encoding genes in the *G. biloba* genome with the protein BLAST (BLASTP) (p -value < 0.001). The significant hits were used as query sequences to search against the National Centre for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/BLAST>) using the translated BLAST (TBLASTN) (p -value < 0.001). The structural integrity of conserved domains was confirmed by Pfam and Simple Modular Architecture Research Tool (SMART, <http://smart.embl-heidelberg.de/>), and redundant sequences were eliminated. The Pfam ID and SMART ID of the domains identified in the DNA methylation pathway gene are detailed in Table S1. The length, molecular weight (MW), and isoelectric point (PI) of confirmed proteins were calculated by ExPasy (https://web.expasy.org/compute_pi/) and the motifs were analyzed by MEME (<http://meme-suite.org/tools/meme>).

2.2. Phylogenetic Analysis and Classification of DNA Methylation Pathway Gene Families in *G. biloba*

The full-length protein sequences of *Arabidopsis thaliana* (TAIR 10 genome release, <https://www.arabidopsis.org/>), *Oryza sativa* (version 5.0, <http://rice.plantbiology.msu.edu/>), *Populus trichocarpa* (version 3.0, <http://www.Phytozome.net>, accessed in September 2006), *Solanum lycopersicum* (version 2.3, <https://solgenomics.net/>, accessed in January 2011), *Pinus tabulaeformis* [37], and the newly identified of *G. biloba* were used to build unrooted phylogenetic trees for MET, CMT, DCL, AGO, and RDR families (Table S2). Sequences of the DNA-methylase domain from various plants were aligned to build unrooted neighbor-joining trees for DRMs and DNMT3s [7]. All DNA methylation pathway protein sequence alignments were made using multiple sequence alignments with high accuracy and high throughput (MUSCLE) algorithm (gap open, −2.9; gap extend, 0; hydrophobicity multiplier, 1.2; clustering method, UPGMB) [8]. Phylogenetic trees for all DNA methylation pathway protein sequences were constructed using the neighbor-joining (NJ) method with the bootstrap test replicated 1000 times, as implemented in MEGA v7.0 [38].

2.3. Chromosomal Localization and Gene Tandem Duplication

Chromosome position and gene tandem duplication were illustrated by TBtools V0.6731 [39]. The tandem duplication of DNA methylation pathway genes was confirmed according to Zhao et al., which included two standards: more than 70% of the longer sequence aligned with the shorter sequence, and a higher than 70% similarity between two aligned sequences [40]. Two genes with five or fewer genes apart in 100-kb chromosome length were considered to be tandem duplicated genes [41].

2.4. Plant Tissues and Abiotic Stress Treatments for the Gene Expression Analysis

To investigate the gene expression patterns in various tissues, the seeds were collected from a 20-year-old ginkgo tree (cv. Jiafoshou), grown at Nanjing Forestry University (32°04′47″ N, 118°48′56″ E), Nanjing, China. After cleaning and removing the outer seed coat, seeds were stratified at 4 °C for 1.5 months. The physiological mature seeds were then germinated at 27 °C and sown into the soil at germinating prophase. Different aged seedlings were then used for both tissue sampling and the abiotic stress treatments. For the young root (YR), young stem (YS), and young leaf (YL) samples, three 1.5-month-old seedlings were collected, of which all roots, main stem without branches, and two uppermost leaves from the same individual seedling were collected, respectively, as one replicate. Mature leaf (ML) and immature fruit (IF) samples were collected from the outer middle part of the mother tree on 2 May, of which each replicate contains two leaves and two fruits, respectively. A total of 15 samples with three replicates for five tissues were all immediately frozen in liquid nitrogen and stored at −80 °C after collecting for future use. Some other tissues were integrated into this

study with their publicly available data, that included: adult root (AR), adult stem (AS), immature leaf (IL, 30 March), microstrobilus (M), ovulate strobilus (OS), and mature fruit (MF, 15 October), with three replicates each, collected from the 31-year-old ginkgo trees (cv. Jiafoshou) planted at Yangtze University, China (30°20′06″ N, 116°19′24″ E) [27]; kernels (with the removal of the testae from the seeds) collected in July (K7), August (K8), September (K9), November (K11), and December (K12) with only one replicate, also from the same mature tree at Nanjing Forestry University [31]; cambium samples from the different aged trees (15Y, 20Y, 22Y, 193Y, 211Y, 236Y, 538Y, 553Y, and 667Y) [30]; and a chichi sample at the elongating stage with a 1.7-cm length collected from Shandong Province, China (35°37′–35°40′ N, 116°36′–117°38′ E) [33].

For the abiotic stress treatments, 35 two-month seedlings, germinating from the seeds described above, were treated with a Hoagland solution with 20% (W/V) PEG6000 to simulate drought stress; two leaves from two random individual seedlings were then collected at 0 (control), 2, 6, 12, and 24 h, respectively, with three replicates. Additionally, 20 six-month seedlings were treated at 40 °C (Heat), with 150 µM NaCl (Salt) and 100 mM methyl jasmonate (MeJA) for 24 h, respectively, then the uppermost leaves from two individual seedlings were collected as one replicate, with three replicates for each; the seedlings before the treatments (0 h) were collected as control. All samples were immediately frozen in liquid nitrogen and stored at −80 °C. The data for the UV-B treatment were downloaded from NCBI (PRJNA595103) [35]. In the experiments, the four-month seedlings were treated with the high-dose UV-B (21.42 kJ/m²/day) for 14 d, then the upper leaves were collected with three replicates; the seedlings treated with the UV-B free white light for 14 d were collected as control.

2.5. Transcriptome Sequencing and Gene Expression Analysis of DNA Methylation Pathway Genes

Total RNAs of collected samples were extracted using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), and then quality-checked using the kaiaoK5500 Spectrophotometer (Kaiao, Beijing, China) and quantified using the RNA Nano 6000 Assay kit (Agilent Technologies, Santa Clara, CA, USA). A total amount of 2 µg RNA per sample was used to generate the sequencing library using NEBNext Ultra RNA Library Prep Kit for Illumina (#E7530, NEB, Ipswich, MA, USA). The libraries were sequenced on an Illumina NovaSeq platform and 150 bp paired-end reads were generated. The raw data have been uploaded to the NCBI database under the project PRJNA650527.

Publicly available transcriptomic data mentioned above for more tissues were downloaded from the NCBI Sequence Read Archive (SRA) database (Table S3). All of the raw RNA-seq data were subject to adapter and low-quality base trimming using Trimmomatic v0.36 [42]. The clean reads were then mapped to the reference genome of *G. biloba* using Spliced Transcripts Alignment to a Reference (STAR) v2.5.3a [43] and the number of reads was calculated for each gene by featureCounts v1.6.2 [44]. The fpkm (fragment counts normalized per kilobase of feature length per million mapped fragments) function of DESeq2 v1.28.1 package was used to calculate and normalize gene expression level using the “median ratio method” [45]. In tissue samples, the Pearson correlation coefficient (PCC) of the data from different replicates was calculated to inspect the consistency (Table S3). The expression value for each gene was represented by the averaged FPKM of the replicates for most of the samples and the FPKM in the only one replicate for the kernel, cambium, and chichi samples. Differentially expressed genes under abiotic stresses were identified using the DESeq2. For each treatment, gene expression in the treated sample was compared with that in the untreated control.

2.6. Coexpression Analysis of DNA Methylation Pathway Genes

RNA-Seq data of 126 *G. biloba* samples, not including the data generated in this study, were downloaded from SRA. Detailed information on these samples is listed in Table S4. The SRA Toolkit v2.9.2 (<https://github.com/ncbi/sra-tools>) was used to convert the raw data into a fastq format. The cleaning, trimming, mapping, and counting of reads were performed as described above. The FPKM values normalized using the “median ratio method” were calculated and applied for co-expression analysis using the weighted gene co-expression network analysis (WGCNA) v1.69 package [46] with

a soft-thresholding power 8. The GO enrichment analysis was performed for genes in each module using the topGO v2.40.0 package [47]. To visualize the connections of 29 genes in the DNA methylation pathway, the Pearson correlation coefficient (PCC) values were extracted and imported into Cytoscape v3.7.1 [48] to generate the network.

3. Results

3.1. Identification and Structural Analysis of DNA Methylation Pathway Gene Families

The hidden Markov model (HMM) profile analyses identified nine DNA methyltransferases (DNMTs) genes in *G. biloba* (Table 1), including two methyltransferases (GbMETs); one chromomethylase (GbCMT); five domains rearranged methyltransferases (GbDRMs); and notably, one GbDNMT3, which was lost in angiosperms and recently functionally identified in *Physcomitrella patens*. Three families involved in the RNA-directed DNA methylation (RdDM) pathway were identified, including four Dicer-likes (GbDCLs), six Argonautes (GbAGOs), and 10 RNA-dependent RNA Polymerases (GbRDRs) (Table 1). The sequences analysis showed that the open reading frames (ORFs) of the GbDNMTs ranged from 665 to 225,575 bp, containing 0 to 22 introns and encoding varied from 144 to 1536 amino acids. The molecular weight (MW) of GbDNMTs ranged from 15.18 to 171.27 KDa, and the protein isoelectric point (PI) ranged from 5.10 to 9.46 (Table 1). Protein structural domain and motif analysis indicated that all of the GbDNMTs have the DNA-methylase domain and most of them have the conserved motif 2, except for three GbDRM2 (Figure 1a, Figure S1). Moreover, GbMETs have two DNMT1-RFD domains and two BAH domains, GbCMT2 has BAH and Chromo domains, and GbDRMs and GbDNMT3 only have the DNA-methylase domain (Figure 1a).

Table 1. Identification and structural analysis of DNA methylation pathway genes.

Gene Family	Gene Name	Gene ID	Chr.	Coordinates (5'–3')	ORF (bp)	Protein (aa)	PI	MW (KDa)	No. of Introns
MET	GbMET1a	Gb_34186	1	204,603,561–204,649,352	45,792	1519	6.21	171.27	9
	GbMET1b	Gb_24436	6	118,058,848–118,066,891	8044	1536	5.78	168.15	8
CMT	GbCMT2	Gb_13672	2	64,359,582–64,366,200	6619	902	8.27	101.47	21
DRM	GbDRMa	Gb_25633	3	282,894,262–282,895,575	1314	437	9.46	49.32	0
	GbDRMb	Gb_33731	12	405,552,612–405,778,186	225,575	650	5.66	73.73	7
	GbDRM2a	Gb_24437	6	118,086,495–118,087,159	665	144	9.33	16.38	2
	GbDRM2b	Gb_03787	11	101,816,084–101,817,173	1090	150	9.18	16.89	3
	GbDRM2c	Gb_23570	7	55,632,811–55,743,991	111,181	202	5.55	22.97	4
DNMT3	GbDNMT3	Gb_23575	9	613,883,922–613,885,240	1319	140	5.10	15.18	3
DCL	GbDCL1	Gb_25982	4	88,761,181–88,828,642	67,462	2164	6.07	244.87	19
	GbDCL2	Gb_32044	5	539,277,719–539,892,110	614,392	1464	6.76	165.28	22
	GbDCL3a	Gb_34767	1	433,398,375–433,857,929	459,555	1815	6.66	203.09	25
	GbDCL3b	Gb_13632	1	428,462,217–428,704,798	459,555	1890	7.57	212.83	25
AGO	GbAGO2a	Gb_13708	9	218,695,767–218,699,026	3260	1025	9.48	113.92	2
	GbAGO2b	Gb_13719	9	219,330,477–219,333,276	2800	896	9.78	100.94	1
	GbAGO3	Gb_21789	9	220,234,425–220,237,707	3283	1025	8.91	115.16	2
	GbAGO7	Gb_15866	11	513,379,788–513,384,603	4816	1199	9.07	134.38	4
	GbAGO10	Gb_34642	1	919,522,451–919,531,642	9192	1061	9.33	119.28	23
	GbAGO18	Gb_34718	11	144,054,917–144,136,109	81,193	741	9.26	83.67	19
RDR	GbRDR1a	Gb_17278	2	146,510,397–146,518,861	8465	915	8.06	103.57	9
	GbRDR1b	Gb_28437	2	147,783,139–147,785,354	2216	618	7.45	69.79	3
	GbRDR1c	Gb_28433	2	149,061,893–149,088,988	27,096	897	8.49	102.07	2
	GbRDR1d	Gb_28429	2	148,784,524–148,786,257	1734	577	9.62	65.37	0
	GbRDR1e	Gb_28426	2	148,423,209–148,425,400	2192	649	9.58	73.35	1
	GbRDR1f	Gb_28427	2	148,451,204–148,455,736	4533	491	6.50	55.78	2
	GbRDR3	Gb_32086	1	607,044,984–607,566,259	521,276	1086	5.88	123.15	18
	GbRDR6a	Gb_12326	3	382,504,304–382,559,266	54,963	1186	8.47	136.02	1
	GbRDR6b	Gb_08298	3	383,759,636–383,761,795	2160	624	8.90	71.62	1
	GbRDR6c	Gb_08306	3	382,920,448–382,945,656	25,209	1187	7.71	135.78	1

Figure 1. Conserved protein domains of different DNA methylation pathway gene families in *G. biloba*. (a) Nine DNA methyltransferases (GbDNMTs), and (b) three families involved in the RNA-directed DNA methylation (RdDM) pathway.

There are four *DCLs* in *G. biloba*, of which two *DCL3* (*GbDCL3a* and *GbDCL3b*) were identified, as previously reported in maize and *Pinus tabulaeformis* [19,37]. *GbDCL1* has the shortest ORF, but with the least number of introns, and encodes a 2164-aa protein, while the other three *GbDCLs* have seven to nine times longer ORFs, but more introns, and encode shorter proteins (Table 1, Figure 1b). The structure of proteins showed seven conserved domains in all four *GbDCLs*, except for an extra Double Stranded RNS-binding (dsRB) domain at the C-terminus of *GbDCL1* (Figure 1b). The same conserved motifs were also observed in all *GbDCLs*, except for *GbDCL1*, which had an extra motif 2 in the middle, and *GbDCL3b*, which had an extra motif 3 (Figure S1).

The six candidate *AGO* genes identified in *G. biloba* have a range from 2800 to 81,193 bp ORFs with one to 23 introns, and encode 741 to 1199-aa proteins (Table 1). Domain analysis using SMART showed that these *GbAGO* proteins contain the same DUF1785, PAZ, and PIWI domains (Figure 1b). The motif analysis also showed that all 10 motifs are conserved in six *GbAGOs*, except for the fact that *GbAGO3* has an extra motif 7 and 10 (Figure S1).

All 10 *RDR* proteins in *G. biloba* share a common RdRP domain corresponding to the catalytic β' subunit of *RDR* [49]. Additionally, an RNA recognition motif (RRM) domain exists at the N-terminus of *GbRDR6a* and *GbRDR6b* (Figure 1b). The *GbRDRs* have a range between 1734 and 521,276 bp ORFs and encode for proteins with 491 to 1187 amino acids. Most of them have zero to three introns, except for *GbRDR3* with 18 introns and *GbRDR1a* with nine introns (Table 1). Interestingly, the RdRP domain of *GbRDR1b* might consist of two isolated segments (95 and 320 amino acids), encoded by two

different portions of the coding region in the *G. biloba* genome, respectively. The motif analysis showed that the GbRDRs can be divided into three sets. The set with GbRDR1a-1e has the conserved motif 1–7; the second set with GbRDR6a and GbRDR6c has all 10 motifs; and the third set with GbRDR1f and RDR3 is special, with only two different motifs (Figure S1).

3.2. Phylogenetic Analysis of DNA Methylation Pathway Genes in *G. biloba*

To study the evolutionary relationship of different DNA methylation pathway gene families, full-length protein sequences from six plant species (one monocot of *Oryza sativa*; three dicots of *Populus trichocarpa*, *Solanum lycopersicum*, and *Arabidopsis thaliana*; and two gymnosperms of *Pinus tabulaeformis* and *G. biloba*) were aligned to construct the unrooted neighbor-joining trees. All DNMTs were clustered into individual subfamilies, in which two GbMETs were close to AtMET1 so were named GbMET1a and GbMET1b, respectively; the only GbCMT was close to OsCMT2 and AtCMT2, so was named GbCMT2 (Figure 2). DNMT3, with a deficiency in angiosperms and overlooked in the past, was identified in this study. The phylogenetic tree showed that the DNMT3s were clustered together with DRMs, which are close homologs in plants, and formed four clades: DRMa/b, DRM2, DRM3, and DNMT3. From three of these clades, two GbDRMs, three GbDRM2s, and one GbDNMT3 were identified, and no DRM3 was obtained in *G. biloba* (Figure 2).

Figure 2. Phylogenetic analysis of DNA methylation pathway gene families. The full-length protein of the monocot of *Oryza sativa*; dicots of *Populus trichocarpa*, *Solanum lycopersicum* and *Arabidopsis thaliana*; and gymnosperms of *Pinus tabulaeformis* were used for tree construction (Table S2). For the domains rearranged methyltransferase (DRM)/DNA methyltransferase 3 (DNMT3) family, the DNA methylation domain region from more species was used [7].

The four GbDCLs were assigned to four groups: DCL1, DCL2, DCL3, and DCL4. All of the GbDCLs were very close to DCLs of *Pinus tabulaeformis*, indicating their conservation among gymnosperms [47]. The DCLs from the clades DCL1 and DCL2 were named GbDCL1 and GbDCL2, respectively. Two proteins identified in the DCL3 clade were very close to PtbDCL3a and PtbDCL3b and were thus named GbDCL3a and GbDCL3b, respectively. The AGO family was clustered into

five clades: AGO1/10, AGO2/3, AGO4, AGO5, and AGO7. GbAGOs were unevenly distributed in the AGO1/10 clade (GbAGO10 and GbAGO18), AGO2/3 clade (GbAGO2a, GbAGO2b, GbAGO3) and AGO7 clade (GbAGO7), with no member in the AGO4 and AGO5 clades. The ten GbRDR proteins were distributed in three clades: RDR1, RDR3, RDR6. Six proteins were aligned to RDR1 and three belonged to the GbRDR6 clade, which were named GbRDR1a/1b/1c/1d/1e/1f and GbRDR6a/6b/6c, respectively. Only one protein was identified in the RDR3 clade, which we named GbRDR3 (Figure 2).

3.3. Chromosomal Location and Tandem Duplication of DNA Methylation Pathway Genes in the *G. biloba* Genome

Nine *GbDNMTs* were found to be distributed on five chromosomes with a variable distribution: *GbMET1a* and *GbMET1b* were located on chromosomes 1 and 6, respectively. The single *GbCMT2* was located on chromosome 2. Each member of *GbDRMs* was positioned separately on five different chromosomes. *GbDNMT3* was located on chromosome 9. Tandem duplication was not found in any DNA methyltransferase family (Figure 3).

Figure 3. Chromosome location of DNA methylation pathway genes in the *G. biloba* genome. Different colors indicate different gene families. Short black curves represent the tandem duplications.

Twenty RdDM related genes were distributed on the seven chromosomes of *G. biloba*. Four *GbDCLs* were located on chromosomes 1, 4, and 5, respectively; six *GbAGOs* were found on chromosomes 1, 9, and 13; and 10 *GbRDRs* were gathered on chromosomes 2 and 3. With more than a 75% similarity of the coding sequence (CDS), tandem duplication was observed for the *GbDCL3a* and *GbDCL3b* on chromosome 1, *GbRDR1e* and *1f* on chromosome 2, and *GbRDR6a* and *6c* on chromosome 3 (Figure 3).

3.4. Expression Patterns of DNA Methylation Pathway Genes in Various Tissues of *G. biloba*

DNA methylation is involved in various aspects of plant growth and development; even sex determination [2,50]. To investigate the expression pattern of DNA methylation pathway genes at different growth and development stages of various tissues, we collected root, stem, leaf, flower, and fruit samples from seedlings or adult plants of *G. biloba* for the transcriptomic analysis. Expressions from some other tissues, including kernels and cambium, were also included using the public dataset. The results showed the different tissue specificities among the genes (Figure 4a). For *GbDNMTs*, *GbDRMs* had a relatively higher expression than *GbMETs* and *GbCMT*. Almost all of the *GbDNMTs*

showed the highest expression in female flower (OS) and fruit (IF and MF) tissues, medium expression in other vegetative tissues (root, stem, and cambium), and low expression in leaves and kernels. *GbDRMb* and *GbDRM2c* were constitutively expressed in all tissues, while *GbDRM2b* was preferentially expressed in the different aged cambium, suggesting their different functions. Furthermore, *GbDNMT3* was hardly detectable in any tissue, although it was identified (Figure 4a).

Figure 4. Expression profile of DNA methylation pathway genes. (a) Expressions in different tissues of *G. biloba*, including the root of young seedlings (YR) and adults (AR); stem of young seedlings (YS) and adults (AS); leaf of young seedlings (YL); immature leaf of adults (IL) and mature leaf (ML); microstrobilus (M) and ovulate strobilus (OS); immature fruit (IF, June) and mature fruit (MF, Oct.); kernels in July (K7), Aug. (K8), Sep. (K9), Nov. (K11) and Dec. (K12); cambium from 15-, 20-, 22-, 193-, 211-, 236-, 538-, 553-, and 667-year-old trees; and chichi (CC). (b) Expression changes under different abiotic stresses. The number shows the log2 ratio of the expression after treatment to the control. Red and blue indicate the genes upregulated (log2 value > 1) and downregulated (log2 value < -1), respectively. NA means not detected in the sample.

Genes involved in the RdDM showed distinct expression patterns. *GbDCL1* was expressed in all tissues, but had a slightly lower expression in the root and stem of seedlings, while *GbDCL2* was preferentially expressed in kernels. Although *GbDCL3a* and *GbDCL3b* are tandem duplicates, they had a higher expression in the cambium, and OS and fruits, respectively, suggesting that their function diverged after the duplication event. Four of the *GbAGOs* showed the highest expression compared with other families. *GbAGO2a* and *GbAGO18* were highly expressed in almost all tissues. *GbAGO10* had a higher expression in the stem, fruits and kernels, and *GbAGO2b* was specifically expressed in the cambium. Most of the *GbRDRs* were highly expressed in the cambium and chichi. *GbRDR1c* and *GbRDR6a* also showed a higher expression in OS and fruits; *GbRDR6c* was preferentially expressed in the adult root (AR), and *GbRDR1f* had a higher expression in mature fruit and early stages of kernels (Figure 4a). These results showed a very differentiated expression pattern between different families and among different members of the same family.

3.5. Expression Changes of DNA Methylation Pathway Genes under Different Abiotic Stresses in *G. biloba*

To study the function of the genes in the abiotic stress responses, we treated the seedlings of *G. biloba* with drought, a high temperature, a high concentration of salt, and methyl jasmonate (MeJA), and the expressions of DNA methylation pathway genes were then detected. The expression under ultraviolet (UV) treatment downloaded from NCBI was also included. The results showed that, under each treatment, at least one member from each family exhibited a changed expression, and several different members responded to different stresses (Figure 4b). Most of the changed genes were involved in multiple stresses, such as *GbCMT2*, *GbDRM2a*, *GbDCL3b*, *GbAGO2b/3*, and *GbRDR1a/1b/1c/6b*, while some others only responded to specific stresses, such as *GbMET1b*, *GbDRMa/2b/2c*, *GbDCL2*, and *GbAGO2a/10*. For the responding *GbDNMTs*, all genes were down-regulated to different extents under different stresses, except for *GbCMT2*, which was significantly up-regulated (by 2.22-fold) under UV treatment. *GbDCL3b* showed an increased expression under high temperature, salt, and MeJA treatment, but a decreased expression under drought. *GbAGO2a* and *GbAGO3* were significantly up-regulated under heat stress, while *GbAGO2b/3/10* was down-regulated under short-time drought stress. *GbAGO2b* was also significantly down-regulated under salt and UV stress. *GbRDRs* were the most differentially regulated under different treatments. Both *GbRDR1b* and *GbRDR1c* were involved in drought stress, but were changed oppositely. *GbRDR1b* also displayed an increased expression under UV stress. However, *GbRDR6b* was significantly down-regulated under salt stress (Figure 4b). In total, more DNA methylation pathway genes were involved under the drought, high temperature, and salt stresses than under MeJA and UV stresses, indicating the different roles of DNA methylation in different abiotic stress responses.

3.6. Coexpression Analysis of DNA Methylation Pathway Genes in *G. biloba*

To explore the function of DNA methylation pathway genes, we downloaded all 135 available public transcriptomic data of *G. biloba* samples, which include varied tissues, mutants, and treatments, from NCBI for the gene coexpression analysis. After filtering out the low-quality data, expressions from 126 samples were used. The expression profiles across all samples showed three main patterns: group I genes (pink) and group II genes (blue) are preferentially expressed in a different subset of samples, respectively, while group III genes are constitutively expressed in all of the samples (Figure 5a). In addition, *GbDNMT3* was not only absent in previous tissues, but also not expressed (FPKM < 1) in any of the samples here, strongly suggesting that it may exist as a pseudogene. Further retrieval of the sample information with high concentrated expressions indicated that group I genes are highly expressed in the ovulate strobilus and fruit, and group II genes are more specifically expressed in the cambium and root (Figure 5a, Table S4). A coexpression analysis was then performed and the network with a Pearson correlation coefficient (PCC) higher than 0.65 was constructed. Consistently, group I and group II genes clustered together, respectively, while group III genes were not closely correlated with each other. In group I, the 11 genes included three families of *GbDNMTs* (*GbMETs*, *GbCMT*, and three *GbDRMs*), two *GbDCLs*, two *GbAGOs*, and one *GbRDR*. In group II, only RdDM-related genes were involved, which included *GbDRM2b*, *GbAGO2b/3*, and six *GbRDRs* (Figure 5b). This suggests that all of the DNA methylation pathway genes may function together as different machines to play a role in different tissues or different growth stages in *G. biloba*.

3.7. Weighted Gene Co-Expression Network Analysis (WGCNA) of DNA Methylation Pathway Genes

The same dataset was then used for the identification of genes coexpressed with DNA methylation pathway genes through weighted gene co-expression network analysis (WGCNA). In total, 18 expressed gene modules were detected, with an average size of 1618 genes, the most genes (4712) in the turquoise module, and the least genes (471) in the grey60 module (Figure 6a).

Figure 5. Coexpression analysis of DNA methylation pathway genes. (a) Expression profile of all genes in the 126 transcriptomic datasets download from NCBI, and (b) coexpression network of DNA methylation pathway genes with a Pearson correlation coefficient (PCC) higher than 0.65. Genes are clustered into two groups (pink and blue).

Relating modules to the expression of DNA methylation pathway genes demonstrated that different modules were correlated with different sets of genes (Figure 6b). The turquoise module was the most highly correlated with group I DNA methylation genes, with most of the PCCs being higher than 0.9, followed by the green, midnightblue and grey60 modules, which were also significantly correlated with the group I genes. The brown module had a close relationship with group II DNA methylation genes, with an average PCC of 0.8, while the blue module showed a significant negative correlation with both groups I and II genes. In addition, some other modules were correlated with other individual DNA methylation genes, such as the salmon module with *GbRDR6c*; pink and salmon modules with *GbAGO2a*; and greenyellow, black, and tan modules with *GbDCL2* (Figure 6b).

To investigate the detailed roles that DNA methylation plays, gene ontology (GO) enrichment analysis was performed for the six highly correlated modules (Figure 7). The results show the very distinguishing enriched biological processes of GO for each module. In the four modules associated with group I, genes of the turquoise module are mainly involved in cell proliferation, genes related to lipid biosynthesis and anion transport are enriched in the midnightblue module, cellular localization-related genes are enriched in the green module, and genes in the grey60 module are involved in polysaccharide metabolism. In the brown module, which is highly correlated with group II DNA methylation genes, signal transduction, cell communication, and the response to stimulus are highly enriched biological processes. In the blue module, which is negatively correlated with almost all of the DNA methylation genes, the photosynthesis and cation transport-related genes are mostly enriched. These results demonstrate that different groups of DNA methylation genes may have differentiated functions, which involve different upstream- or downstream- genes.

Figure 6. Weighted gene co-expression network analysis (WGCNA) of DNA methylation pathway genes. (a) Expressed genes were assigned to 18 modules, and (b) the correlation between each module and individual DNA methylation genes. Pink line and blue line indicate group I and group II genes, respectively.

3.8. Key Genes Coexpressed with DNA Methylation Pathway Genes

Focusing on the whole DNA methylation pathway, instead of the individual methylation genes, we identified the top 30 genes that coexpressed with the whole group I or group II DNA methylation genes, respectively (Figure 8, Table S5). Reasonably, the group I coexpressed genes all belong to the turquoise module from the WGCNA. Consistent with the GO enrichment results, it includes the genes directly involved in DNA replication and the cell cycle, such as *replication factor C subunit 2*, *DNA helicase*, *double strand RNA binding protein*, *chromatin assembly factor*, *kinetochore protein*, *chromatin remodeling 24*, and *histone H2AXa*. It also has some genes encoding transcription factors, such as the NAC domain protein, receptor protein kinase-like protein, and protease with unknown functions. The group II coexpressed genes all fall into the brown module, including coding genes of RNA-helicase, E3 ubiquitin-protein ligase, protein kinases, transcription factors, and more unknown proteins, which may function in cell communication and the response to stimulus (Figure 8). These genes provide the candidates that the DNA methylation pathway genes may interact with or regulate directly for downstream functions.

Figure 7. Gene Ontology (GO) enrichment of genes from six coexpressed WGCNA modules. p values were calculated using topGO and the top10 enriched GO terms in biological processes were listed.

Figure 8. Top 30 genes coexpressed with group I (pink) and group II (blue) DNA methylation pathway genes, respectively. Turquoise and brown show the modules that the coexpressed genes belong to. The cut-off Pearson Correlation Coefficient (PCC) of 0.85 was used.

4. Discussion

4.1. The Structure of DNA Methylation Pathway Genes

In this study, four subfamilies of DNMTs were identified in *G. biloba*, among which METs, CMTs, and DRMs are commonly present in plants, while the fourth one—DNMT3—is absent in most angiosperms and had only been functionally identified in the gymnosperm, *Physcomitrella patens*. It has been reported that DNMT3 mediates the de novo methylation of CG and CHG in a way that is independent from DRM [7]. However, the expression of *GbDNMT3* was barely detectable in various tissues and all of the public data, suggesting that *GbDNMT3* may be a pseudogene and has lost its function in *G. biloba*. Moreover, most angiosperms have DNMT2 [51–53], but we did not find it in *G. biloba*. The function of DNMT2 may be replaced by DRM. *GbMETs* are conserved between *G. biloba* and *Arabidopsis*, and all have two DNA-methylase domains [51].

Only three classes of DCL genes have been identified in *G. biloba*, including *GbDCL1*, *GbDCL2*, and *GbDCL3* (containing *GbDCL3a* and *GbDCL3b*), which is different from a previous study, which revealed four families containing one member each [54]. This was possibly because of the updated genome version we used in this study. Moreover, different DCLs are responsible for the formation of specific small RNAs. DCL1 plays an essential role in the formation of microRNA (miRNA). DCL2 is responsible for producing 22-nt siRNA, and DCL4 is related to the generation of trans-acting small interfering RNAs (ta-siRNAs) and 21-nt siRNAs [55]. DCL3 generates 24-nt siRNAs which participate in RdDM [56,57]. However, the functions of various DCLs in *G. biloba* still need to be studied. Based on protein sequence analysis, the identity of *GbDCL3a* and *GbDCL3b* is 54.39%, while the identity of *GbDCL3a* and *PtcDCL3a* is 59.63%, and the identity of *GbDCL3b* and *PtcDCL3b* is 60.16%. This suggests that *GbDCL3a* and *GbDCL3b* were separated before the divergence of the two species, consistent with the results obtained for *Pinus tabulaeformis* [37].

The AGO family is a highly basic binding protein, which is characterized by PAZ and PIWI domains [58]. All kinds of sRNAs produced by DCLs combine with the AGO protein to form the core of RNA-induced silencing complexes (RISCs) [55]. Six AGOs were identified in *G. biloba*, including *GbAGO2a*, *2b*, *3*, *7*, *10*, and *18*. Interestingly, *AGO1* and *AGO4* were identified in most of the plants, while not obtained in our study. Moreover, *GbAGO2a*, *GbAGO2b*, and *GbAGO3* are close to each other in the phylogenetic tree, and their positions on the chromosomes are also close, which is consistent with previous studies on *Arabidopsis*. However, the results show that the function of *AtAGO3* is different from that of *AtAGO2* and it mainly recruits 24-nt siRNA to regulate epigenetic silencing [59].

The function of RDRs is to synthesize the precursor of double-stranded RNA with the single-stranded RNA template [18]. There are two types of RDRs in *Arabidopsis thaliana*, including RDR α (*AtRDR1*, *AtRDR2* and *AtRDR6*), which has been well-studied, and RDR γ (*AtRDR3*, *AtRDR4* and *AtRDR5*), which still has an unknown function [60,61]. The major function of *AtRDR1* is an anti-virus function. Using the single-stranded RNA produced by the virus, *AtRDR2* is related to the synthesis of endogenous siRNAs (hsRNAs), and *AtRDR6* is involved in the formation of ta-siRNAs and nat-siRNAs [60]. In the present study, the *G. biloba* genome encoded three types of RDR genes, including *RDR1*, *RDR3* and *RDR6*. *GbRDR3* is the only acidic protein of the RDR family and is located on chromosome 1 alone, while the function of its homolog in *Arabidopsis* is still unclear.

4.2. The Expression Pattern of DNA Methylation Pathway Genes in Various Tissues and under Abiotic Stresses

The transcript abundance of methylation pathway genes is uneven in different gene families and family members. It has been reported that DNA methylation plays an important role in reproductive development and sex determination [62]. In *Pinus tabulaeformis*, most of the DNA methylation-related sRNA pathway genes have higher expression levels in female than in male cones [37]. In *Populus tomentosa*, DNA methylation genes *MET1* and *DECREASED DNA METHYLATION 1 (DDM1)* are located in a sex-determination region in the sex chromosome, and the expression levels in female flowers are higher than in male flowers [50]. In maize, sex determination

is associated with DNA methylation genes [63]. In the present study, half of the DNA methylation pathway genes, especially the group I genes, had higher expression levels in the ovulate strobilus (female flower) than in the microstrobilus (male flower), consistent with a previous study demonstrating that the differential expression levels of DNA methylation genes may be related to sex determination in *G. biloba* [62]. Some genes also had higher expression levels in fruit, suggesting that DNA methylation may be involved in reproductive development. In addition, the RdDM-related genes are more highly expressed in meristem tissues than in tissues that grow primarily by cell expansion in *Arabidopsis thaliana* [64]. In the present study, most of the genes also had high expression levels in the cambium compared with the other vegetative tissues, which is consistent with the study on *Arabidopsis thaliana*. This indicates that DNA methylation may play an indispensable role in the vegetative growth of *G. biloba*. Furthermore, *GbRDR1b* and *GbRDR1f* and *GbRDR6a* and *GbRDR6c* are both tandem duplication pairs with similar expression patterns, while *GbDCL3a* and *GbDCL3b* have quite different expressions in various tissues and under stresses, suggesting the divergence of their functions after duplication in *G. biloba*.

Most of the DNA methylation pathway genes (15/29) in this study showed a changed expression under different abiotic stresses. *GbAGO2a* was significantly up-regulated (5.25-fold) under heat stress, which is consistent with the change after two-hour treatment in maize. *GbAGO2b* exhibited a decreased expression under salt and drought treatments, similar to the expression changes in maize [65]. *GbRDR1a* was significantly up-regulated after six-hour drought treatment, which was congruent with the result obtained in tomato [18]. However, *GbDRM2a* was down-regulated under salt stress, while *DRM2* showed an increased expression in chickpea roots [66]. Similarly, the expression of *GbDCL3* was decreased (0.32-fold after 6h and 0.47-fold after 12 h) under drought stress, while it was increased in pepper [58]. These results indicate that some gene family members are functionally conserved across species, while other members play roles in different ways. Moreover, certain *G. biloba* genes (such as *GbCMT2*, *GbDCL3b*, and *GbAGO3*) have distinct expression patterns under different abiotic stresses, suggesting that they might respond to different abiotic stresses with specific mechanisms [67].

4.3. The Differentiated Functions of DNA Methylation Pathway Genes

The coexpression network of DNA methylation pathway genes showed that they were clustered into two groups. Group I contains three types of *GbDNMTs* and three other RdDM-related families, which covers the genes responsible for the de novo methylation and maintenance of all three types of DNA methylation (CG, CHG, and CHH). All of these genes were preferentially expressed in the ovulate strobilus and fruits in *G. biloba*. The genes of group II are only RdDM-related and more specifically expressed in the cambium. WGCNA analysis revealed that the genes of group I and group II are highly correlated with different modules, in which distinct GO enrichments were observed, e.g., group I genes were closely related to the turquoise module where cell proliferation genes were enriched, while the cell communication and signal transduction genes were enriched in the brown module, which significantly relates to group II genes. This indicates that multi-member families in the same pathway may differentiate into separate machines to work independently, with same or different functions, in the entire or specific tissues, respectively. This agrees with a previous report on rice, illustrating that 27 starch synthesis genes of seven families were clustered into two groups, and were responsible for starch biosynthesis in the leaf (source) and seed (sink), respectively [68].

It should be noted that there were no *DCLs* in group II genes. Previous studies have shown that two-thirds of the target region of RdDM still existed even when most of the siRNA disappeared in *dcl* mutants, signifying that there is a DCL-independent RdDM pathway mediated by some DCL-independent siRNA in *Arabidopsis thaliana* [11,69]. Our results imply that a DCL-independent RdDM pathway may play a role in the growth and development of *G. biloba*, especially in the cambium. However, this needs to be further confirmed by experiments.

5. Conclusions

In this study, we systematically identified DNA methylation pathway genes by bioinformatics methods, and analyzed their potential functions by WGCNA in *G. biloba*. Twenty-nine genes of seven families were identified and clustered into two groups according to their expression. Through WGCNA, two modules (turquoise and brown modules) demonstrated the highest correlation with group I and group II DNA methylation genes, respectively, in which the genes encoding enzymes or transcription factor, such as Gb_08699, Gb_08989 and Gb_15278 in the turquoise module, were believed to play an important role in DNA repair or cell proliferation, while the genes in the brown module, such as the transcription factors Gb_07584 and Gb_06045, may be very important for adapting to environmental stress. In addition, most of the DNA methylation pathway genes were highly expressed in reproductive tissues and meristem tissues, consistent with the previous research in rice and *Arabidopsis thaliana* [10,64]. In summary, we identified the DNA methylation pathway genes and explored their expression pattern in various tissues and environmental conditions, and further identified the genes highly correlated with them, laying a foundation for DNA methylation related research in *G. biloba*. A further DNA methylation level test would be helpful for better understanding its epigenetic effects. With the development of research methods, DNA methylation-related gene editing, the modification of specific methylation sites, or developing epimarkers could greatly boost molecular design breeding or forest selection and cultivation.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4907/11/10/1076/s1>, Figure S1: Motif and gene structure of DNA methylation genes in *Ginkgo biloba*; Figure S2: Visualization of the eigengene network representing the relationships among the modules and the DNA methylation pathway genes; Figure S3: Gene Ontology (GO) enrichment of genes from six correlated WGCNA modules. *p* values were calculated using topGO; Table S1: Pfam and SMART domain IDs of DNA methylation pathway proteins; Table S2: Gene ID of *MET*, *CMT*, *AGO*, *RDR* and *DCL* gene families used for the phylogenetic analysis; Table S3: RNA-seq data of different tissue samples collected from this study and SRA database; Table S4: RNA-seq data of 126 samples downloaded from NCBI SRA database and used for WGCNA; Table S5: Top 30 coexpressed genes with two groups of DNA methylation pathway genes.

Author Contributions: Conceptualization, F.-F.F. and C.G.; Data curation, J.C., Y.S., Z.Z. and T.Z.; Formal analysis, C.G., M.D. and X.Y.; Investigation, W.Y.; Resources, W.Y., Y.S., Z.Z. and T.Z.; Supervision, F.-F.F., L.X., F.C. and G.W.; Writing—original draft, C.G.; Writing—review and editing, F.-F.F., X.Y., J.C., L.X., F.C. and G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 31971689.

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

References

1. Dinh, T.T.; O’Leary, M.; Won, S.Y.; Li, S.; Arroyo, L.; Liu, X.; Defries, A.; Zheng, B.; Cutler, S.R.; Chen, X. Generation of a luciferase-based reporter for CHH and CG DNA methylation in *Arabidopsis thaliana*. *Silence* **2013**, *4*, 1. [CrossRef] [PubMed]
2. Fan, S.; Liu, H.; Liu, J.; Hua, W.; Xu, S.; Li, J. Systematic Analysis of the DNA Methylase and Demethylase Gene Families in Rapeseed (*Brassica napus* L.) and Their Expression Variations After Salt and Heat stresses. *Int. J. Mol. Sci.* **2020**, *21*, 953. [CrossRef] [PubMed]
3. Zhu, C.; Zhang, S.; Zhou, C.; Chen, L.; Fu, H.; Li, X.; Lin, Y.; Lai, Z.; Guo, Y. Genome-wide investigation and transcriptional analysis of cytosine-5 DNA methyltransferase and DNA demethylase gene families in tea plant (*Camellia sinensis*) under abiotic stress and withering processing. *PEERJ* **2020**, *8*, e8432. [CrossRef] [PubMed]
4. Stroud, H.; Do, T.; Du, J.; Zhong, X.; Feng, S.; Johnson, L.; Patel, D.J.; Jacobsen, S.E. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat. Struct. Mol. Biol.* **2014**, *21*, 64–72. [CrossRef] [PubMed]
5. Bewick, A.J.; Schmitz, R.J. Gene body DNA methylation in plants. *Curr. Opin. Plant Biol.* **2017**, *36*, 103–110. [CrossRef] [PubMed]
6. Arkan, B.; Özden, S.; Turgut-Kara, N. DNA methylation related gene expression and morphophysiological response to abiotic stresses in *Arabidopsis thaliana*. *Environ. Exp. Bot.* **2018**, *149*, 17–26. [CrossRef]

7. Yaari, R.; Katz, A.; Domb, K.; Harris, K.D.; Zemach, A.; Ohad, N. RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. *Nat. Commun.* **2019**, *10*, 1613. [[CrossRef](#)] [[PubMed](#)]
8. Sabbione, A.; Daurelio, L.; Vegetti, A.; Talón, M.; Tadeo, F.; Dotto, M. 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.* **2019**, *19*, 1–13. [[CrossRef](#)]
9. Yadav, C.B.; Muthamilarasan, M.; Pandey, G.; Prasad, M. Identification, Characterization and Expression Profiling of Dicer-Like, Argonaute and RNA-Dependent RNA Polymerase Gene Families in Foxtail Millet. *Plant Mol. Biol. Rep.* **2015**, *33*, 43–55. [[CrossRef](#)]
10. Kapoor, M.; Arora, R.; Lama, T.; Nijhawan, A.; Khurana, J.P.; Tyagi, A.K.; Kapoor, S. Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genom.* **2008**, *9*, 451. [[CrossRef](#)]
11. Zhang, H.; Lang, Z.; Zhu, J.K. Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 489–506. [[CrossRef](#)] [[PubMed](#)]
12. Jones, A.L.; Sung, S. Mechanisms Underlying Epigenetic Regulation in *Arabidopsis thaliana*. *Integr. Comp. Biol.* **2014**, *54*, 61–67. [[CrossRef](#)] [[PubMed](#)]
13. Huang, H.; Liu, R.; Niu, Q.; Tang, K.; Zhang, B.; Zhang, H.; Chen, K.; Zhu, J.; Lang, Z. Global increase in DNA methylation during orange fruit development and ripening. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1430–1436. [[CrossRef](#)] [[PubMed](#)]
14. Daccord, N.; Celton, J.; Linsmith, G.; Becker, C.; Choise, N.; Schijlen, E.; van de Geest, H.; Bianco, L.; Micheletti, D.; Velasco, R.; et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* **2017**, *49*, 1099–1106. [[CrossRef](#)]
15. Bitonti, M.B.; Cozza, R.; Chiappetta, A.; Giannino, D.; Castiglione, M.R.; Dewitte, W.; Mariotti, D.; Onckelen, H.V.; Innocenti, A.M. Distinct nuclear organization, DNA methylation pattern and cytokinin distribution mark juvenile, juvenile-like and adult vegetative apical meristems in peach (*Prunus persica* (L.) Batsch). *J. Exp. Bot.* **2002**, *53*, 1047–1054. [[CrossRef](#)]
16. Palmgren, G.; Mattsson, O.; Okkels, F.T. Specific Levels of DNA Methylation in Various Tissues, Cell Lines, and Cell Types of *Daucus carota*. *Plant Physiol.* **1991**, *95*, 174–178. [[CrossRef](#)]
17. Bartels, A.; Han, Q.; Nair, P.; Stacey, L.; Gaynier, H.; Mosley, M.; Huang, Q.; Pearson, J.; Hsieh, T.; An, Y.; et al. Dynamic DNA Methylation in Plant Growth and Development. *Int. J. Mol. Sci.* **2018**, *19*, 2144. [[CrossRef](#)]
18. Bai, M.; Yang, G.; Chen, W.; Mao, Z.; Kang, H.; Chen, G.; Yang, Y.; Xie, B. 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* **2012**, *501*, 52–62. [[CrossRef](#)]
19. Qian, Y.; Cheng, Y.; Cheng, X.; Jiang, H.; Zhu, S.; Cheng, B. Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Rep.* **2011**, *30*, 1347–1363. [[CrossRef](#)]
20. Zhao, K.; Zhao, H.; Chen, Z.; Feng, L.; Ren, J.; Cai, R.; Xiang, Y. The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: Gene structure, gene expression, phylogenetic analysis and evolution. *J. Genet.* **2015**, *94*, 317–321. [[CrossRef](#)]
21. Das, S.; Meher, P.K.; Rai, A.; Bhar, L.M.; Mandal, B.N. Statistical Approaches for Gene Selection, Hub Gene Identification and Module Interaction in Gene Co-Expression Network Analysis: An Application to Aluminum Stress in Soybean (*Glycine max* L.). *PLoS ONE* **2017**, *12*, e169605. [[CrossRef](#)] [[PubMed](#)]
22. Childs, K.L.; Davidson, R.M.; Buell, C.R. Gene coexpression network analysis as a source of functional annotation for rice genes. *PLoS ONE* **2011**, *6*, e22196. [[CrossRef](#)] [[PubMed](#)]
23. Weston, D.J.; Karve, A.A.; Gunter, L.E.; Jawdy, S.S.; Yang, X.; Allen, S.M.; Wulschleger, S.D. Comparative physiology and transcriptional networks underlying the heat shock response in *Populus trichocarpa*, *Arabidopsis thaliana* and *Glycine max*. *Plant Cell Environ.* **2011**, *34*, 1488–1506. [[CrossRef](#)]
24. Downs, G.S.; Bi, Y.; Colasanti, J.; Wu, W.; Chen, X.; Zhu, T.; Rothstein, S.J.; Lukens, L.N. A Developmental Transcriptional Network for Maize Defines Coexpression Modules. *Plant Physiol.* **2013**, *161*, 1830–1843. [[CrossRef](#)]
25. Ficklin, S.P.; Feltus, F.A. Gene Coexpression Network Alignment and Conservation of Gene Modules between Two Grass Species: Maize and Rice. *Plant Physiol.* **2011**, *156*, 1244–1256. [[CrossRef](#)] [[PubMed](#)]

26. Sabater-Jara, A.B.; Souliman-Youssef, S.; Novo-Uzal, E.; Almagro, L.; Belchí-Navarro, S.; Pedreño, M.A. Biotechnological approaches to enhance the biosynthesis of ginkgolides and bilobalide in *Ginkgo biloba*. *Phytochem. Rev.* **2013**, *12*, 191–205. [[CrossRef](#)]
27. Ye, J.; Cheng, S.; Zhou, X.; Chen, Z.; Kim, S.U.; Tan, J.; Zheng, J.; Xu, F.; Zhang, W.; Liao, Y.; et al. A global survey of full-length transcriptome of *Ginkgo biloba* reveals transcript variants involved in flavonoid biosynthesis. *Ind. Crop. Prod.* **2019**, *139*, 111547. [[CrossRef](#)]
28. Guo, J.; Wu, Y.; Wang, G.; Wang, T.; Cao, F. Integrated analysis of the transcriptome and metabolome in young and mature leaves of *Ginkgo biloba* L. *Ind. Crop. Prod.* **2020**, *143*, 111906. [[CrossRef](#)]
29. Guo, J.; Zhou, X.; Wang, T.; Wang, G.; Cao, F. Regulation of flavonoid metabolism in ginkgo leaves in response to different day-night temperature combinations. *Plant Physiol. Biochem.* **2020**, *147*, 133–140. [[CrossRef](#)]
30. Wang, L.; Cui, J.; Jin, B.; Zhao, J.; Xu, H.; Lu, Z.; Li, W.; Li, X.; Li, L.; Liang, E.; et al. Multifeature analyses of vascular cambial cells reveal longevity mechanisms in old *Ginkgo biloba* trees. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 2201–2210. [[CrossRef](#)]
31. He, B.; Gu, Y.; Xu, M.; Wang, J.; Cao, F.; Xu, L.A. Transcriptome analysis of *Ginkgo biloba* kernels. *Front. Plant Sci.* **2015**, *6*, 819. [[CrossRef](#)] [[PubMed](#)]
32. Li, D.; Wu, D.; Li, S.; Guo, N.; Gao, J.; Sun, X.; Cai, Y. Transcriptomic profiling identifies differentially expressed genes associated with programmed cell death of nucellar cells in *Ginkgo biloba* L. *BMC Plant Biol.* **2019**, *19*, 91. [[CrossRef](#)] [[PubMed](#)]
33. Liu, X.; Sun, L.; Wu, Q.; Men, X.; Yao, L.; Xing, S. Transcriptome profile analysis reveals the ontogenesis of rooted chichi in *Ginkgo biloba* L. *Gene* **2018**, *669*, 8–14. [[CrossRef](#)] [[PubMed](#)]
34. Roodt, D.; Lohaus, R.; Sterck, L.; Swanepoel, R.L.; van de Peer, Y.; Mizrachi, E. Evidence for an ancient whole genome duplication in the cycad lineage. *PLoS ONE* **2017**, *12*, e184454. [[CrossRef](#)] [[PubMed](#)]
35. Zhao, B.; Wang, L.; Pang, S.; Jia, Z.; Wang, L.; Li, W.; Jin, B. UV-B promotes flavonoid synthesis in *Ginkgo biloba* leaves. *Ind. Crop. Prod.* **2020**, *151*, 112483. [[CrossRef](#)]
36. Ma, X.; Zhao, H.; Xu, W.; You, Q.; Yan, H.; Gao, Z.; Su, Z. Co-expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development. *Front. Genet.* **2018**, *9*, 574. [[CrossRef](#)]
37. Niu, S.; Liu, C.; Yuan, H.; Li, P.; Li, Y.; Li, W. Identification and expression profiles of sRNAs and their biogenesis and action-related genes in male and female cones of *Pinus tabulaeformis*. *BMC Genom.* **2015**, *16*, 693. [[CrossRef](#)]
38. Zhao, H.; Zhao, K.; Wang, J.; Chen, X.; Chen, Z.; Cai, R.; Xiang, Y. Comprehensive Analysis of Dicer-Like, Argonaute, and RNA-dependent RNA Polymerase Gene Families in Grapevine (*Vitis Vinifera*). *J. Plant Growth Regul.* **2015**, *34*, 108–121. [[CrossRef](#)]
39. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
40. Zhao, P.; Wang, D.; Wang, R.; Kong, N.; Zhang, C.; Yang, C.; Wu, W.; Ma, H.; Chen, Q. Genome-wide analysis of the potato Hsp20 gene family: Identification, genomic organization and expression profiles in response to heat stress. *BMC Genom.* **2018**, *19*, 61. [[CrossRef](#)]
41. Wang, L.; Guo, K.; Li, Y.; Tu, Y.; Hu, H.; Wang, B.; Cui, X.; Peng, L. Expression profiling and integrative analysis of the CESA/CSL superfamily in rice. *BMC Plant Biol.* **2010**, *10*, 282. [[CrossRef](#)] [[PubMed](#)]
42. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
43. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)] [[PubMed](#)]
44. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [[CrossRef](#)]
45. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
46. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)]
47. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017; Available online: <http://www.R-project.org/> (accessed on 25 September 2020).

48. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
49. Iyer, L.M.; Koonin, E.V.; Aravind, L. Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct. Biol.* **2003**, *3*, 1. [[CrossRef](#)]
50. Song, Y.; Ma, K.; Ci, D.; Chen, Q.; Tian, J.; Zhang, D. Sexual dimorphic floral development in dioecious plants revealed by transcriptome, phytohormone, and DNA methylation analysis in *Populus tomentosa*. *Plant Mol. Biol.* **2013**, *83*, 559–576. [[CrossRef](#)]
51. Pavlopoulou, A.; Kossida, S. Plant cytosine-5 DNA methyltransferases: Structure, function, and molecular evolution. *Genomics* **2007**, *90*, 530–541. [[CrossRef](#)]
52. Ahmad, F.; Huang, X.; Lan, H.X.; Huma, T.; Bao, Y.M.; Huang, J.; Zhang, H.S. Comprehensive gene expression analysis of the DNA (cytosine-5) methyltransferase family in rice (*Oryza sativa* L.). *Genet. Mol. Res.* **2014**, *13*, 5159–5172. [[CrossRef](#)] [[PubMed](#)]
53. Qian, Y.; Xi, Y.; Cheng, B.; Zhu, S. Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize. *Plant Cell Rep.* **2014**, *33*, 1661–1672. [[CrossRef](#)] [[PubMed](#)]
54. Ma, L.; Hatlen, A.; Kelly, L.J.; Becher, H.; Wang, W.; Kovarik, A.; Leitch, I.J.; Leitch, A.R. Angiosperms Are Unique among Land Plant Lineages in the Occurrence of Key Genes in the RNA-Directed DNA Methylation (RdDM) Pathway. *Genome Biol. Evol.* **2015**, *7*, 2648–2662. [[CrossRef](#)] [[PubMed](#)]
55. Fang, X.; Qi, Y. RNAi in Plants: An Argonaute-Centered View. *Plant Cell* **2016**, *28*, 272–285. [[CrossRef](#)] [[PubMed](#)]
56. Qin, H.; Chen, F.; Huan, X.; Machida, S.; Song, J.; Yuan, Y.A. Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA* **2010**, *16*, 474–481. [[CrossRef](#)] [[PubMed](#)]
57. Moura, M.O.; Fausto, A.K.S.; Fanelli, A.; Guedes, F.A.D.F.; Silva, T.D.F.; Romanel, E.; Vaslin, M.F.S. 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 Biol.* **2019**, *19*, 503. [[CrossRef](#)]
58. Qin, L.; Mo, N.; Muhammad, T.; Liang, Y. Genome-Wide Analysis of DCL, AGO, and RDR Gene Families in Pepper (*Capsicum Annuum* L.). *Int. J. Mol. Sci.* **2018**, *19*, 1038. [[CrossRef](#)]
59. Zhang, Z.; Liu, X.; Guo, X.; Wang, X.; Zhang, X. Arabidopsis AGO3 predominantly recruits 24-nt small RNAs to regulate epigenetic silencing. *Nat. Plants* **2016**, *2*, 16049. [[CrossRef](#)]
60. Polydore, S.; Axtell, M.J. Analysis of RDR1/RDR2/RDR6-independent small RNAs in *Arabidopsis thaliana* improves MIRNA annotations and reveals unexplained types of short interfering RNA loci. *Plant J.* **2018**, *94*, 1051–1063. [[CrossRef](#)]
61. Willmann, M.R.; Endres, M.W.; Cook, R.T.; Gregory, B.D. The Functions of RNA-Dependent RNA Polymerases in *Arabidopsis*. *Arab. Book* **2011**, *9*, e146. [[CrossRef](#)]
62. Du, S.; Sang, Y.; Liu, X.; Xing, S.; Li, J.; Tang, H.; Sun, L. Transcriptome Profile Analysis from Different Sex Types of *Ginkgo biloba* L. *Front. Plant Sci.* **2016**, *7*, 871. [[CrossRef](#)] [[PubMed](#)]
63. Parkinson, S.E.; Gross, S.M.; Hollick, J.B. Maize sex determination and abaxial leaf fates are canalized by a factor that maintains repressed epigenetic states. *Dev. Biol.* **2007**, *308*, 462–473. [[CrossRef](#)]
64. Baubec, T.; Finke, A.; Mittelsten, S.O.; Pecinka, A. Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. *EMBO Rep.* **2014**, *15*, 446–452. [[CrossRef](#)] [[PubMed](#)]
65. Zhai, L.; Teng, F.; Zheng, K.; Xiao, J.; Deng, W.; Sun, W. Expression analysis of Argonaute genes in maize (*Zea mays* L.) in response to abiotic stress. *Hereditas* **2019**, *156*, 27. [[CrossRef](#)] [[PubMed](#)]
66. Garg, R.; Kumari, R.; Tiwari, S.; Goyal, S. Genomic survey, gene expression analysis and structural modeling suggest diverse roles of DNA methyltransferases in legumes. *PLoS ONE* **2014**, *9*, e88947. [[CrossRef](#)] [[PubMed](#)]
67. Xie, Z.; Johansen, L.K.; Gustafson, A.M.; Kasschau, K.D.; Lellis, A.D.; Zilberman, D.; Jacobsen, S.E.; Carrington, J.C. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2004**, *2*, E104. [[CrossRef](#)] [[PubMed](#)]

68. Fu, F.F.; Xue, H.W. Coexpression analysis identifies Rice Starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol.* **2010**, *154*, 927–938. [[CrossRef](#)]
69. Yang, D.L.; Zhang, G.; Tang, K.; Li, J.; Yang, L.; Huang, H.; Zhang, H.; Zhu, J.K. Dicer-independent RNA-directed DNA methylation in *Arabidopsis*. *Cell Res.* **2016**, *26*, 66–82. [[CrossRef](#)]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).