



Article

High-Depth Transcriptome Reveals Differences in Natural Haploid *Ginkgo biloba* L. Due to the Effect of Reduced Gene Dosage

Yaping Hu ^{1,2} , Petr Šmarda ², Ganping Liu ¹, Beibei Wang ¹, Xiaoge Gao ¹ and Qirong Guo ^{1,*}

¹ Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China

² Department of Botany and Zoology, Masaryk University, Koltlářská 2, 61137 Brno, Czech Republic

* Correspondence: qrguo@njfu.edu.cn

Abstract: As a representative of gymnosperms, the discovery of natural haploids of *Ginkgo biloba* L. has opened a new door for its research. Haploid germplasm has always been a research material of interest to researchers because of its special characteristics. However, we do not yet know the special features and mechanisms of haploid ginkgo following this significant discovery. In this study, we conducted a homogenous garden experiment on haploid and diploid ginkgo to explore the differences in growth, physiology and biochemistry between the two. Additionally, a high-depth transcriptome database of both was established to reveal their transcriptional differences. The results showed that haploid ginkgo exhibited weaker growth potential, lower photosynthesis and flavonoid accumulation capacity. Although the up-regulated expression of DEGs in haploid ginkgo reached 46.7% of the total DEGs in the whole transcriptome data, the gene sets of photosynthesis metabolic, glycolysis/gluconeogenesis and flavonoid biosynthesis pathways, which were significantly related to these differences, were found to show a significant down-regulated expression trend by gene set enrichment analysis (GSEA). We further found that the major metabolic pathways in the haploid ginkgo transcriptional database were down-regulated in expression compared to the diploid. This study reveals for the first time the phenotypic, growth and physiological differences in haploid ginkgos, and demonstrates their transcriptional patterns based on high-depth transcriptomic data, laying the foundation for subsequent in-depth studies of haploid ginkgos.

Keywords: ginkgo; haploid; transcriptome; gene dosage; mechanism



Citation: Hu, Y.; Šmarda, P.; Liu, G.; Wang, B.; Gao, X.; Guo, Q. High-Depth Transcriptome Reveals Differences in Natural Haploid *Ginkgo biloba* L. Due to the Effect of Reduced Gene Dosage. *Int. J. Mol. Sci.* **2022**, *23*, 8958. <https://doi.org/10.3390/ijms23168958>

Academic Editor: Frank M. You

Received: 2 June 2022

Accepted: 10 August 2022

Published: 11 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In general, high-ploidy plants of the same species tend to exhibit enhanced growth vigor compared to low ploidy ones [1]. This enhanced vigor is reflected by a larger body size [2], faster growth rate [3], and more secondary metabolites [4]. In crops, polyploid technology has been used to develop more advantageous varieties of *Cucumis melo* L. [5], tetraploid hybrid rice [6], decaploid strawberries [7], and so on. Researchers have mainly focused on polyploids and neglected the special value of haploids, as haploid plants are generally dwarfed and have low growth. However, haploids have incomparable advantages in shortening the breeding cycle and accelerating the breeding process [8,9].

Numerous diploid plant species contain spontaneous haploids. However, such haploids appear to be uncommon in gymnosperms [10,11]. Since the exploration of natural haploid plants in *Datura stramonium* in 1922 [12], haploids are becoming a prominent study focus in plant germplasm research. Haploid rarely exists in nature as a species. Almost all the extant haploids are created by the techniques such as anther culture, parthenogenesis, etc. [13]. Pure line haploid germplasm resources are of great significance for breeding improvement and accelerating the process of developing improved varieties. The discovery of haploids has greatly helped in deciphering highly heterozygous plant

genomes [14,15]. Therefore, an increasing number of researchers are working on the induction of haploids [16–18]. The regeneration system of ginkgo has not yet been established, and various methods of inducing haploids in vitro are also effective in ginkgo [19].

Ginkgo biloba L. originated in the Carboniferous period, then underwent quaternary glaciation, and finally survived only in China. From the 18th century, ginkgo has spread from China to Europe via Japan and Korea [20,21]. Ginkgo is widely used in daily production and life as an important economic tree species. For instance, flavonoids play an important role in *G. biloba* extract (GBE), which is widely employed in clinical treatment and is one of the main sources of commercial value for ginkgo [22,23]. Ginkgo flavonoid content is affected by many factors, such as fertilization [24], salt [25], temperature [26], light [27], etc. Many regulatory genes and transcription factors (TFs) have been identified to resolve the regulatory network of ginkgo flavonoids [28–30]. However, all these studies were based on diploid ginkgo, as no one had found ginkgo of another ploidy until 2016. We first discovered natural ginkgo haploids using flow cytometry and investigated their cytological properties in 2016 [31,32]. The discovery of ginkgo haploids shattered the earlier perception that ginkgo is only diploid. Exploring the differences in haploid compared to diploid ginkgo is important for understanding the effects of gene dosage on gene expressions.

In this study, we obtained high-depth transcriptomic data of natural haploid and diploid ginkgos and used these data to reveal the underlying mechanisms of many differences in haploid ginkgos. This was the first report of transcriptomic differences following the discovery of natural haploid ginkgo material.

2. Results

2.1. Haploid Ginkgo Exhibits Weak Growth, Biomass Accumulation and Secondary Metabolic Synthesis

Just as most polyploid plants tend to exhibit relatively high growth potential, haploid ginkgos are weaker than diploids in growth, photosynthesis, and secondary metabolite accumulation (Figure 1). The haploid ginkgo leaf area ($7.9 \pm 0.5 \text{ cm}^2$) was significantly smaller than that of the diploid ($17.6 \pm 1.1 \text{ cm}^2$) when the functional leaves matured in August (Figure 1). Similarly, the net photosynthetic rate and flavonoid concentration of haploid ginkgo were substantially lower than that of diploids ($p < 0.05$). The net photosynthetic rate of haploid ginkgo was 36.1% that of diploid. The lack of photosynthetic capacity leads to weaker biomass accumulation in haploid ginkgo morphology. We used flow cytometry (FLC) to identify the ploidy of the haploid and diploid ginkgo cultivars (Supplementary File S1).

2.2. Transcriptomic Analysis Overview

To further explain the differences in transcriptional regulation among different ploidy ginkgos, we selected haploid ‘Rocky’ and diploid ‘Maronia’ with large differences in flavonoid content for transcriptome analysis. Six RNA libraries were generated and analyzed (H1, H2, and H3 for haploid ginkgo, and D1, D2, and D3 for diploid ginkgo). RNA-Seq generated 149.82 GB of clean data (Q30 > 94%) with an average of 24.97 Gb per sample (Table 1). In order to explore the transcriptional differences between haploid and diploid ginkgo to the greatest extent, the amount of transcriptome data in this study was much larger than that of conventional transcriptomes. There were 486.71 million high-quality clean readings collected in total, ranging from 73.31 to 89.43 million per sample. The mapped clean reads were assembled into 29,036 genes (Supplementary File S2) after alignment with the latest ginkgo reference genome. Among these genes, a total of 2452 genes (Supplementary File S3) were identified as differentially expressed (DEGs), with 1149 upregulated and 1303 downregulated (haploid vs. diploid).

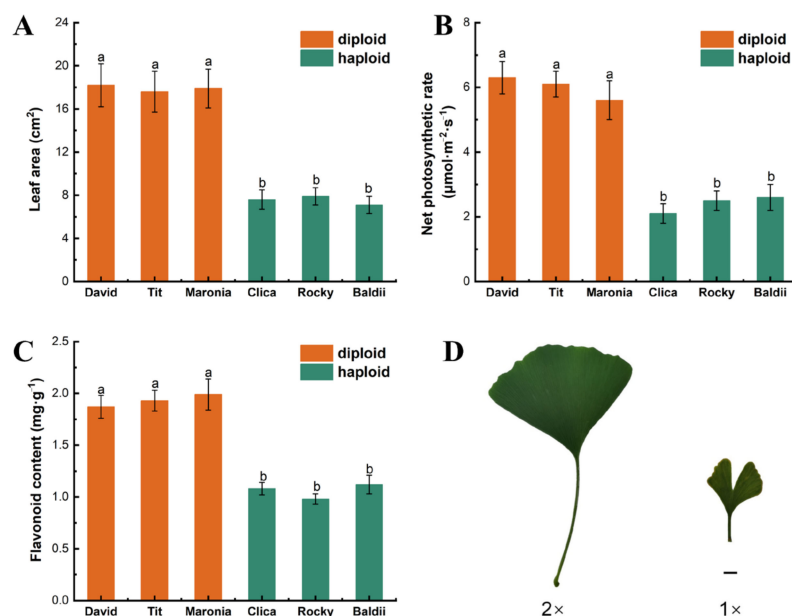


Figure 1. Comparison of morphology, growth, photosynthesis and flavonoid content of haploids and diploids. Schemes follow the same formatting. **(A)** Leaf area. **(B)** The net photosynthetic rate of the leaves. **(C)** Leaf flavonoid content. **(D)** Haploid and diploid leaf morphology. Bars with different letters indicate significant differences at $p < 0.05$ according to Duncan's test; 1×: haploid plants; 2×: diploid plants. Bar = 1 cm.

Table 1. An overview of the sequencing quality of six ginkgo RNA libraries.

Sample	Clean Reads	Clean Data (bp)	Q20 (%)	Q30 (%)	GC (%)
H1	84,053,070	12,606,818,190	97.84	94.66	44.72
H2	74,996,246	10,743,266,079	97.76	94.37	44.65
H3	73,313,140	10,496,771,635	97.87	94.66	44.55
D1	89,428,662	13,348,479,805	97.80	94.55	44.39
D2	82,482,584	12,311,759,601	97.79	94.46	44.48
D3	82,336,494	12,322,466,518	97.91	94.79	44.47

2.3. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Analysis of DEGs

Overall, 34 biological process (BP) phrases, 9 cellular component (CC) terms, and 35 molecular function (MF) terms were found to be enriched in 2485 DEGs. Among the top 20 GO-terms of enrichment, 'detoxification', 'cell wall organization', and 'external encapsulating structure organization' were the most enriched in the BP categories (Figure 2A). 'Peroxidase activity', 'oxidoreductase activity, acting on peroxide as acceptor', and 'ADP binding' were the top three subcategories of enrichment degree in the MF category. In the CC category, the subcategory with the highest enrichment degree was 'cell wall', followed by 'external encapsulating structure' and 'extracellular region'.

KEGG analysis of transcriptomics showed that 502 candidate genes with pathway annotations were enriched in 106 KEGG pathways (Supplementary File S4). 'Monoterpenoid biosynthesis' was the most substantially enriched pathway, with five up-regulated and three down-regulated genes (Figure 2B), followed by 'Flavone and flavonol biosynthesis' and 'Cutin, suberin and wax biosynthesis'.

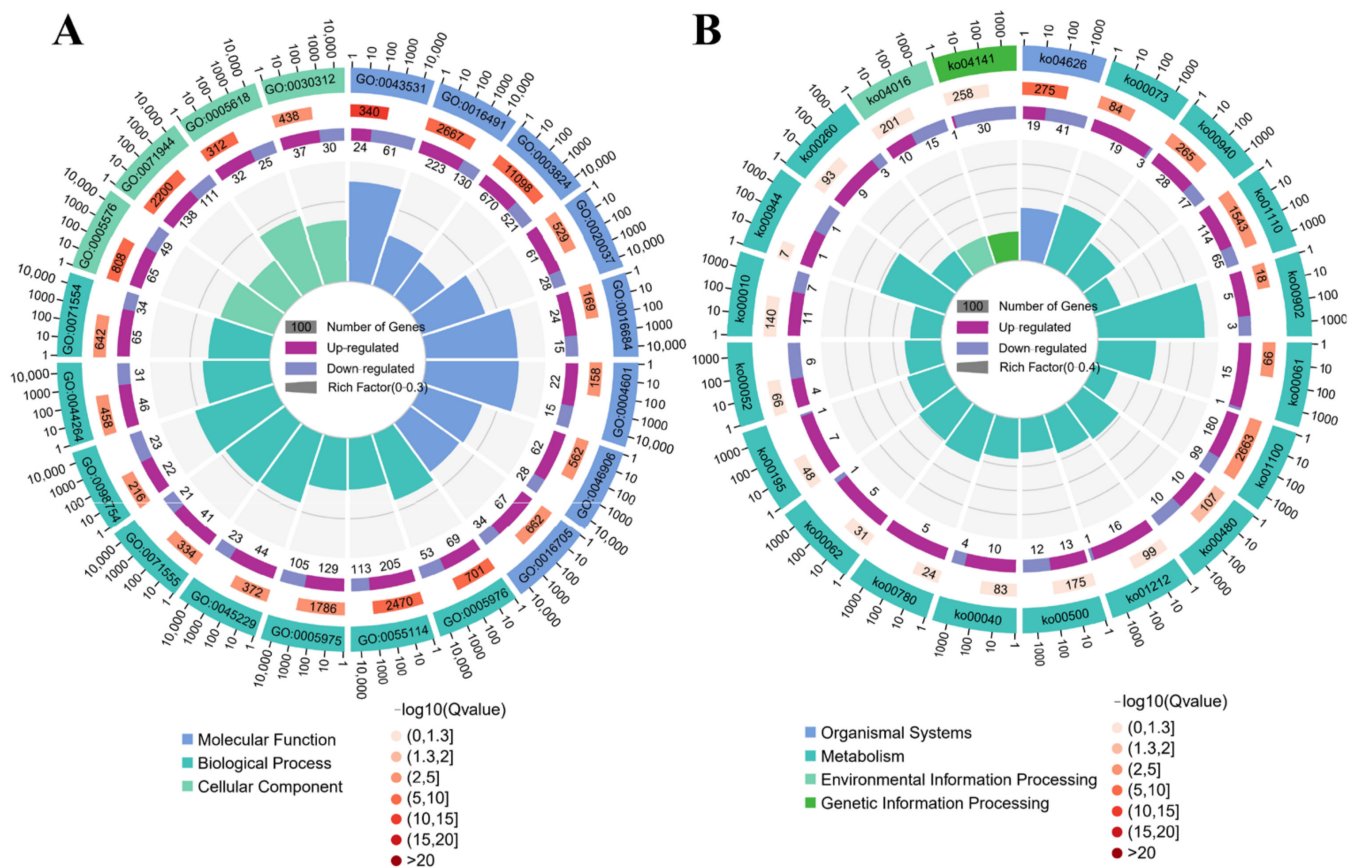


Figure 2. DEG enrichment analysis utilizing GO and KEGG pathways. **(A)** GO classifications of DEGs. **(B)** KEGG classifications of DEGs. The first circle from outside and inside is the pathway of the top 20 enrichment terms, and the number outside the circle is the gene number scale. The second circle represents the number and Q value of the pathway in the background gene. The third circle represents the number and ratio of up- (light purple) and down-regulated (dark purple) genes. The fourth circle represents the Rich Factor value of each pathway. GO:0043531, ADP binding. GO:0005976, polysaccharide metabolic process. GO:0016491, oxidoreductase activity. GO:0003824, catalytic activity. GO:0003824, catalytic activity. GO:0005618, cell wall. GO:0071944, cell periphery. GO:0055114, oxidation-reduction process. GO:0030312, external encapsulating structure. GO:0005975, carbohydrate metabolic process. GO:0016684, oxidoreductase activity, acting on peroxide as acceptor. GO:0004601, peroxidase activity. GO:0004601, peroxidase activity. GO:0045229, external encapsulating structure organization. GO:0020037, heme binding. GO:0044264, cellular polysaccharide metabolic process. GO:0071554, cell wall organization or biogenesis. GO:0098754, detoxification. GO:0006633, fatty acid biosynthetic process. GO:0016798, hydrolase activity, acting on glycosyl bonds. ko04626, Plant-pathogen interaction. ko00073, Cutin, suberine and wax biosynthesis. ko00940, Phenylpropanoid biosynthesis. ko01110, Biosynthesis of secondary metabolites. ko00061, Fatty acid biosynthesis. ko00902, Monoterpenoid biosynthesis. ko01100, Metabolic pathways. ko00480, Glutathione metabolism. ko00480, Glutathione metabolism. ko00500, Starch and sucrose metabolism. ko00040, Pentose and glucuronate interconversions. ko00062, Fatty acid elongation. ko00062, Fatty acid elongation. ko00780, Biotin metabolism. ko00195, Photosynthesis. ko00052, Galactose metabolism. ko00010, Glycolysis/Gluconeogenesis. ko04016, MAPK signaling pathway-plant. ko00944, Flavone and flavonol biosynthesis. ko00130, Ubiquinone and other terpenoid-quinone biosynthesis.

2.4. Differential Analysis of Photosynthesis Metabolic Pathway

The net photosynthesis plays a crucial role in plant growth and development. Compared to diploids, haploid ginkgo has seven DEGs enriched in the photosynthesis pathway, six of which were down-regulated and one up-regulated in expression (Figure 3). F-type ATPase, widely found in chloroplast vesicle-like membranes, is known as the world's smallest and most sophisticated molecular motor [33]. Two key genes regulating F-type ATPase activity, *atpF* and *atpA*, were significantly down-regulated in haploid ginkgo. Additionally, in the reaction process of photosystem II, *rpoA* and *PSBW* were also significantly down-regulated. However, not all genes in photosystem II were down-regulated in haploid ginkgo, such as *petH*. Based on these analyses, we tentatively hypothesized that the weaker photosynthetic capacity of haploid ginkgo was possibly due to the down-regulation of some key genes in the photosynthetic response process.

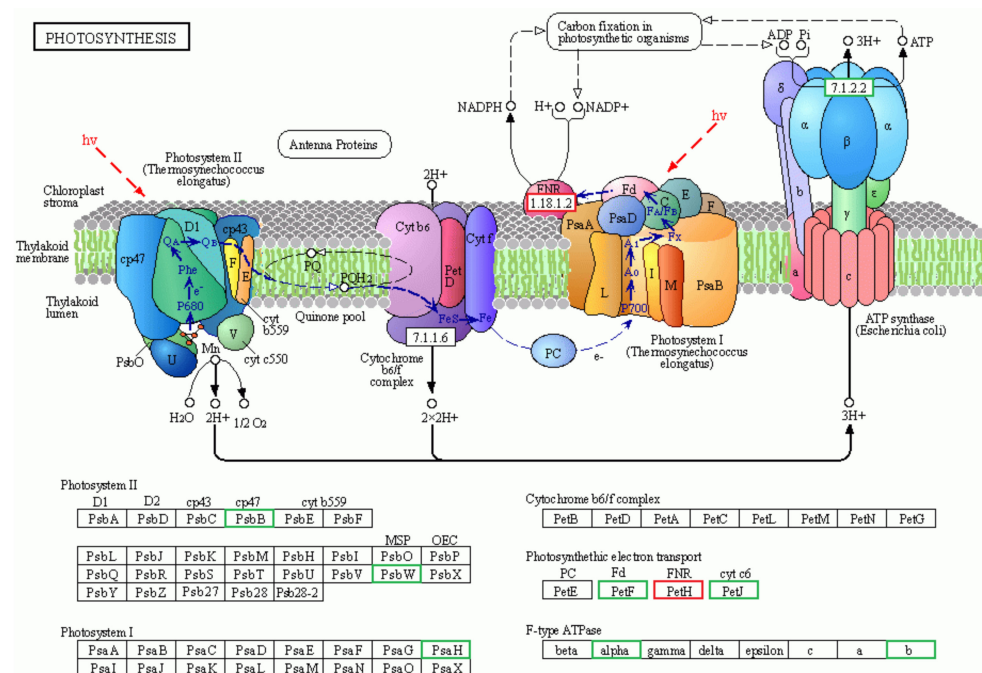


Figure 3. Schematic diagram of DEGs in the photosynthesis pathway of haploid ginkgo. Green indicated down-regulated expression and red indicated up-regulated expression.

Enrichment analysis based on traditional hypergeometric tests often requires the use of significantly different gene set data. When changes in the expression of individual genes are weak, little or no results may be obtained based on traditional enrichment analysis. Gene set enrichment analysis (GSEA) can effectively make up for the lack of effective information for minor genes in traditional enrichment analysis, and more comprehensively explain the regulatory role of a functional unit [34,35]. We set up all genes in the photosynthesis pathway as a gene set to analyze the overall expression of photosynthesis genes in haploid ginkgo (Figure 4). The defined core set of regulatory genes was shown in Supplementary File S5. We could find that the expression of genes related to plant photosynthesis showed a significant decrease in haploid ginkgo. Therefore, we can clearly state that the lower photosynthesis capacity of haploid ginkgo was due to the low dosage effect of photosynthesis-related genes.

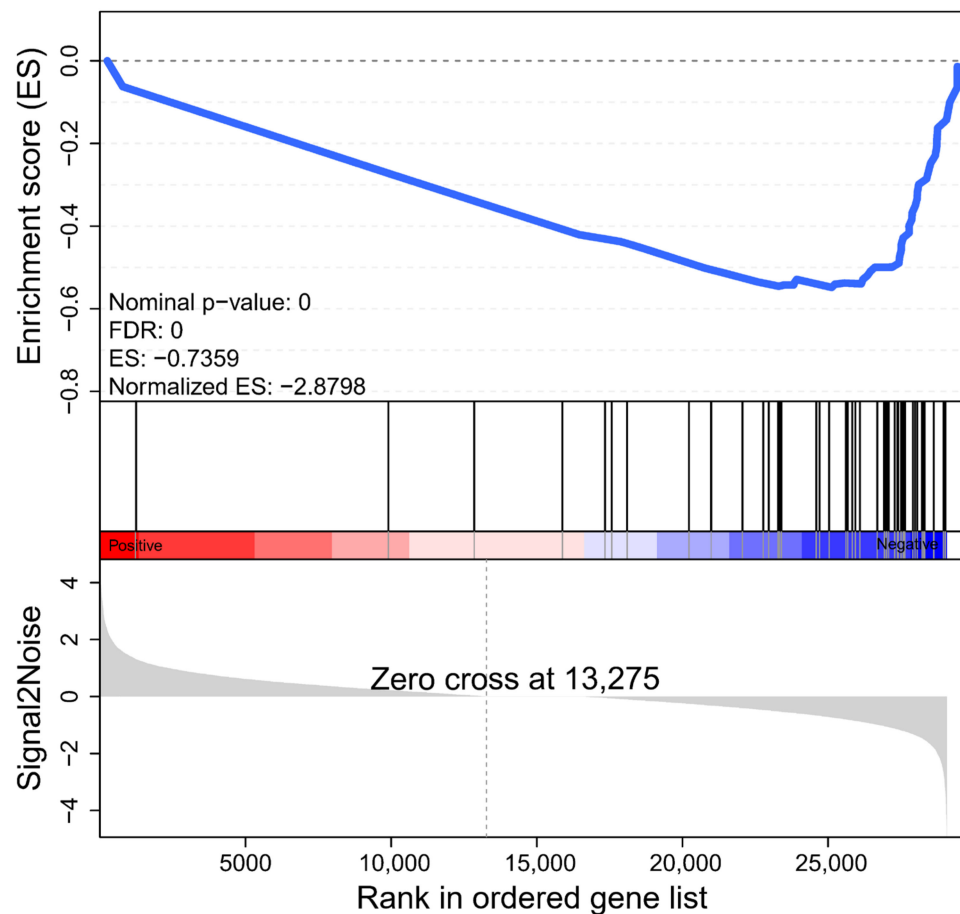


Figure 4. GSEA of all genes in the photosynthesis pathway. The folded line is the enrichment score (ES) of each gene, and the peak value is the ES of the gene set. The horizontal axis represents each gene in this gene set and corresponds to a vertical line similar to a barcode. ES: Enrichment Score. Normalized ES: Normalized ES value after correction. FDR: p -value after correction for multiple hypothesis testing. Probability estimates of possible false positive results for Normalized ES. The smaller the FDR, the more significant the enrichment is. Core genes are the genes before the peak for gene sets with positive ES. The core genes are the genes after the peak for the gene set with negative ES. Overall, if genes are enriched at the top of the curve, the gene set is up-regulated expression trend, and conversely, if they are enriched at the bottom, the expression trend is down-regulated.

2.5. Differential Analysis of Glycolysis/Gluconeogenesis Pathway

The primary physiological function of sugars is to provide the body with the energy it needs for its vital activities. Gluconeogenesis is the main way in which organisms obtain energy and is a common stage through which all organisms must undergo gluconeogenesis. A total of 17 DEGs were enriched in the glycolysis/gluconeogenesis pathway of haploid ginkgo, of which 11 were up-regulated and 6 were down-regulated in expression (Figure 5). In terms of the number of DEGs in the glycolysis/gluconeogenesis metabolic pathway alone, the number of up-regulated DEGs expressed in haploid ginkgo was even greater than the number of down-regulated expression. However, it was found by GSEA that the entire set of genes regulating glycolysis/gluconeogenesis metabolism containing 140 genes showed a pattern of reduced expression in haploid ginkgo (Figure 6). The entire gene set can be found in Supplementary File S6.

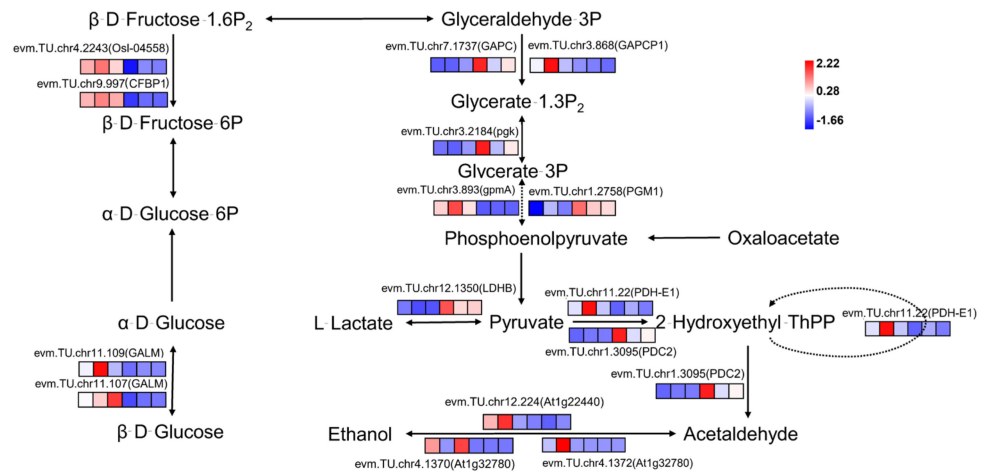


Figure 5. Schematic diagram of DEGs in the glycolysis/gluconeogenesis pathway of haploid ginkgo. The three squares on the left of the gene expression heatmap represent haploid ginkgo, and the three squares on the right represent diploid ginkgo.

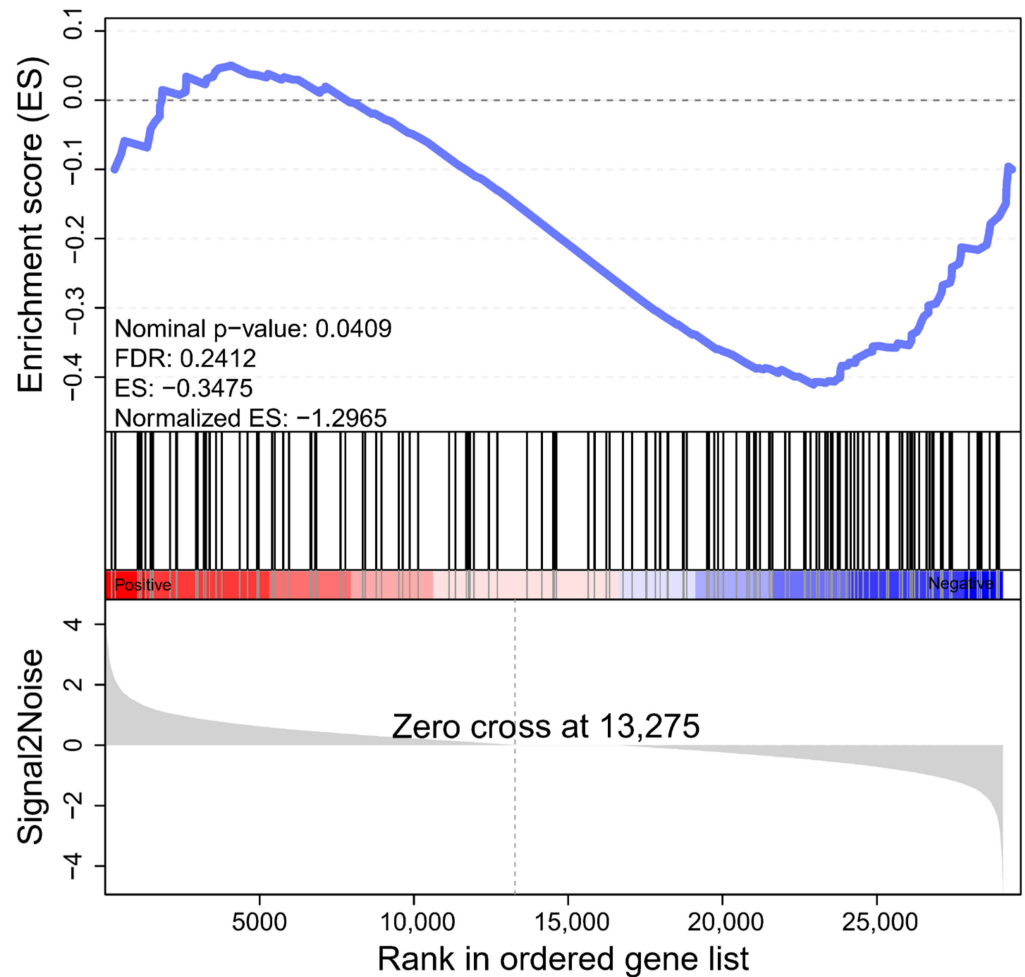


Figure 6. GSEA of all genes in the glycolysis/gluconeogenesis pathway.

2.6. Differential Analysis of Flavonoid Biosynthesis Pathway

Just as haploid ginkgos exhibit lower photosynthesis capacity and growth potential, haploids also had lower flavonoid content than diploids. Unlike the model plants where the flavonoid biosynthesis pathways were relatively well understood, the regulatory network of flavonoid biosynthesis in ginkgo is not yet fully defined. Some key regulatory genes in the flavonoid biosynthesis pathway of haploid ginkgo were down-regulated in expression, such as *F3H*, *FLS*, *HST*, and *CYP75A* (Figure 7). However, the key genes regulating the production of 5-Deoxyleucocyanidin, Leucopelargonidin, and Luteoforol, *DFR*, was up-regulated in haploid ginkgo. In addition, the key gene, *LAR*, which regulates Afzelechin and (+)-Gallocatechin production, was also down-regulated in expression.

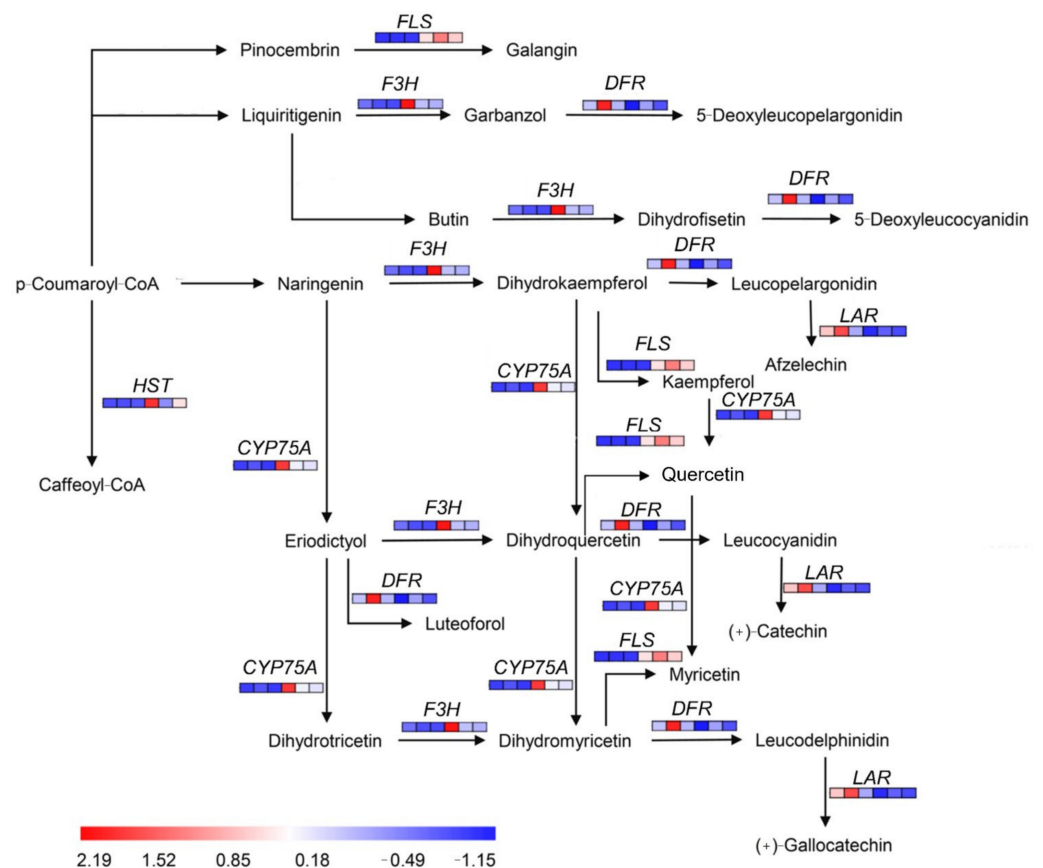


Figure 7. Schematic diagram of DEGs in the flavonoid biosynthesis pathway of haploid ginkgo.

Although the expression of DEGs in the haploid ginkgo flavonoid biosynthesis pathway was inconsistent, we could see by GSEA that the genes related to flavonoid biosynthesis showed an overall trend of down-regulated expression (Figure 8). The defined core set of regulatory genes was shown in Supplementary File S7. However, it was still possible to find that there are some genes that do not follow this trend. This also indicates that there are still unexplored corners of the ginkgo flavonoid regulatory network.

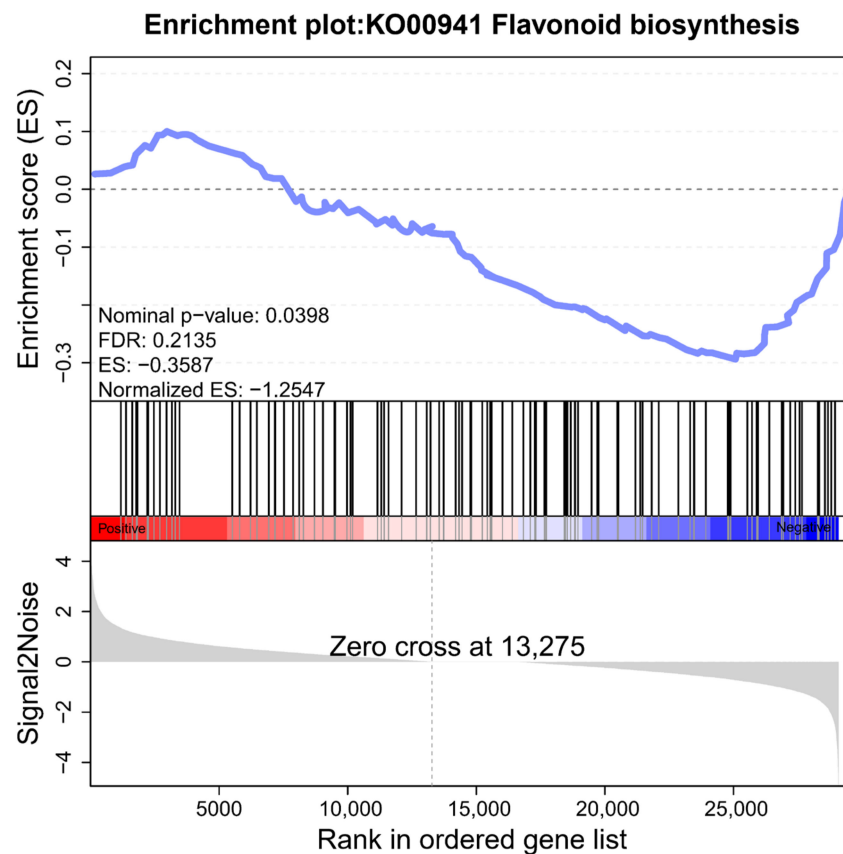


Figure 8. GSEA of all genes in the flavonoid biosynthesis pathway.

2.7. Enrichment Analysis of Up-Regulated DEGs in Haploid Ginkgo

The up-regulated genes in haploid ginkgo accounted for 46.7% of the entire differentially expressed gene set. We performed GO enrichment analysis on the up-regulated genes in haploids and found that the most significant enriched term for these genes was ‘diaminobutyrate acetyltransferase activity’ (Figure 9A). The most significantly enriched KEGG pathway of up-regulated DEGs was ‘Plant-pathogen interaction’, followed by ‘Protein processing in endoplasmic reticulum’ and ‘Glutathione metabolism’ (Figure 9B). GSEA results on up-regulated expression of DEGs in haploid ginkgo showed that only two metabolic pathways were significantly enriched, ‘Protein processing in endoplasmic reticulum’ and ‘MAPK signaling pathway—plant’ (Figure 10).

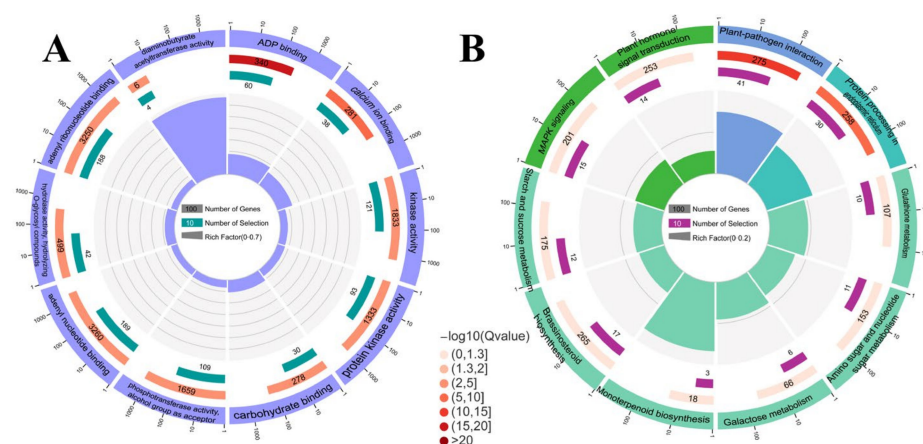


Figure 9. The up-regulated DEGs enrichment analysis utilizing GO (A) and KEGG (B) pathways.

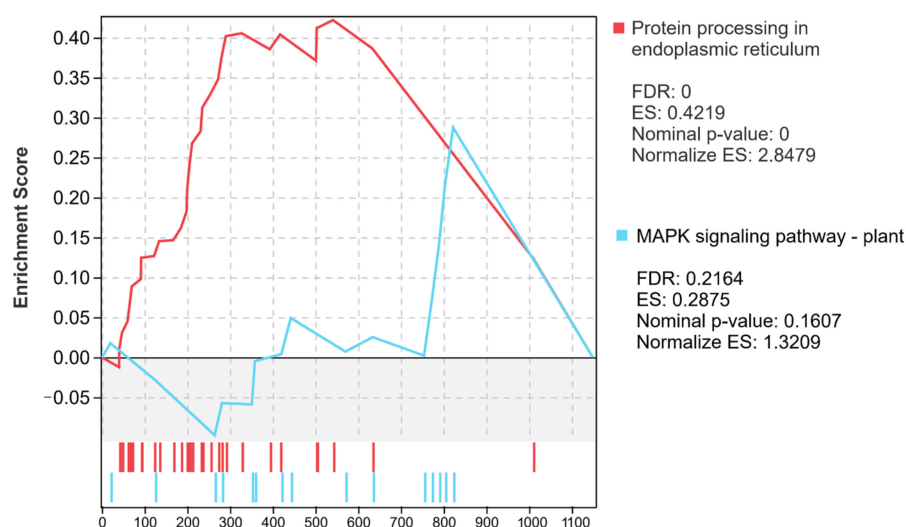


Figure 10. GSEA of up-regulated DEGs in haploid ginkgo.

2.8. An Overall Trend of Down-Regulated Expression of DEGs in Haploid Ginkgo

We further performed GSEA on all the major enriched pathways of DEGs in the whole transcriptome database, and the results all remained consistent. A total of 21 KEGG metabolic pathways were significantly enriched and showed a trend of down-regulated expression (Figure 11, Supplementary File S8). The top three significantly enriched pathways were ‘Photosynthesis’, ‘Fatty acid biosynthesis’, and ‘Cutin, suberine and wax biosynthesis’.

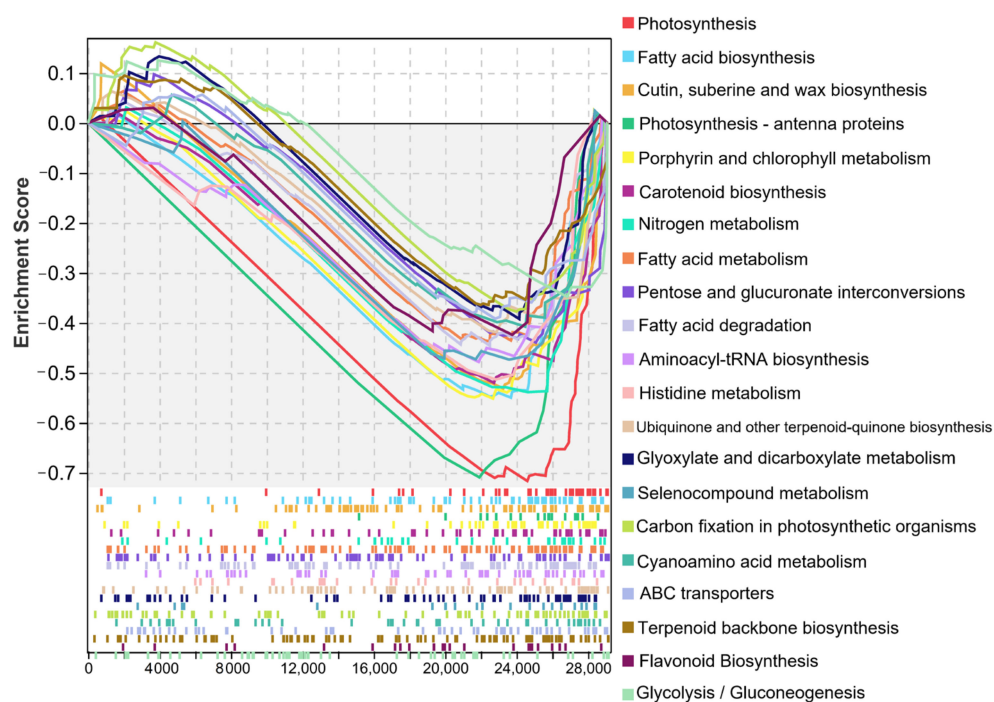


Figure 11. The metabolic pathways significantly enriched in GSEA.

2.9. Transcript Level and Quantitative Real-Time PCR (qRT-PCR) Validation of mRNA in Haploid and Diploid Ginkgo

Fourteen randomly selected DEGs in the ‘Flavonoid biosynthesis’ and ‘Phenylpropanoid biosynthesis’ pathways were used to perform qRT-PCR to verify that the transcription group data is valid and reliable (Figure 12). These genes’ real-time fluorescence expression trends were consistent with the transcriptome data, showing that our transcriptome data was credible. Significantly, the expression of several genes in the ‘Phenylpropanoid

biosynthesis' pathway was significantly higher in haploid ginkgos than in diploids, such as *evm.TU.chr4.2086*, *evm.TU.chr10.1202* and *evm.TU.chr10.109*. The 'Phenylpropanoid biosynthesis' pathway, as an important upstream regulatory pathway of 'Flavonoid biosynthesis', is one of our priorities and the next step to be taken.

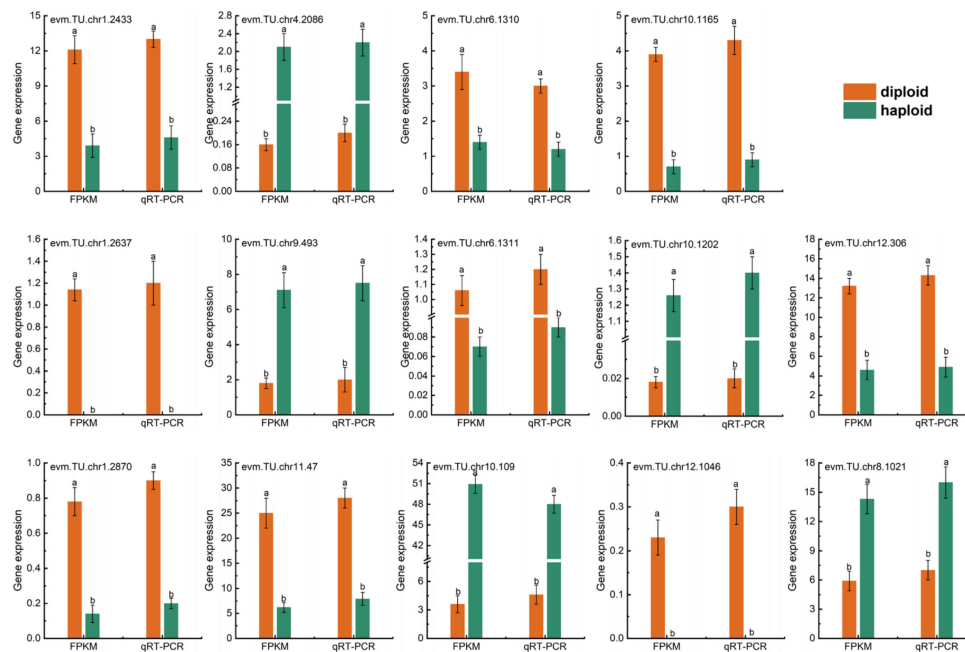


Figure 12. Validation of 14 chosen genes by qRT-PCR in haploid and diploid ginkgo leaves. Bars with different letters indicate significant differences at $p < 0.05$ according to Duncan's test.

3. Discussion

As with most species, polyploid plants are taller and grow faster, and haploid plants are relatively short and weak, as is the case with ginkgo. Polyploid ginkgos have larger stomata and cells [32]. Larger cells and more cellular content can accumulate more energy for their growth and development. Qualitative differences in morphology, growth, and physiology and biochemistry among ginkgos of different ploidy follow the laws of ploidy, as has been demonstrated on *Liriodendron sino-americanum* [36], *Eucommia* [37], rice [38], *Rhododendron fortunei* [39], and *Thymus vulgaris* [40]. However, this is not absolute; for example, the size of pure tetraploid apples (*Malus × domestica*) and *Citrus limonia* are smaller compared to their diploid counterparts [41,42].

The increasing amount of gene expression data in ploidy studies has stimulated scientists to speculate on the regulatory mechanisms driving these gene expression changes. Plant polyploidization may cause DNA sequence mutations such as chromosomal rearrangements and gene loss, DNA modifications, and changes in gene expression [43,44]. In the present study, haploid ginkgos showed different degrees of DEGs compared to diploid ginkgos. We focused on metabolic pathways that exhibit apparent differences in photosynthesis, glycolysis/gluconeogenesis, and flavonoid biosynthesis. Although the extent of changes in the expression of DEGs in these pathways was not entirely consistent, the GSEA showed an overall trend of down-regulation in haploid ginkgo. Haploid ginkgos present a marked reduction in gene dosage compared to diploids. Studies have confirmed that DNA methylation and small RNAs play an important role in the expression of different genes [45].

However, within this overall trend of down-regulated expression, up-regulated expression of genes in haploid ginkgo accounted for 46.7% of the total differentially expressed genes. However, these DEGs were only significantly enriched in the 'Protein processing in endoplasmic reticulum' and 'MAPK signaling pathway-plant' pathways by GSEA. 'MAPK signaling pathway-plant' is widely involved in plant immunity and environmental stress

response. [46]. Haploids are less stable in the natural environment and less resistant to adversity than diploids. Genes in the ‘MAPK signaling pathway–plant’ pathway were up-regulated in haploid ginkgo, signaling that haploid ginkgo was trying to adapt to the habitat.

The traits of an organism are regulated by genes, and transcription is an important but not all-inclusive part of the transmission of genetic information. Proteins are the executors of genetic information. Combining the proteomics of haploid ginkgo may better explain the reasons for these differences.

We are more interested in haploid ginkgo than in other ploidy types, as haploids are perfect materials for perfecting the highly heterozygous, oversized mega-genome of ginkgo [47]. Our next step is to further improve the quality of the ginkgo reference genome by using haploid material. The use of haplotype genomes to resolve complex trait regulatory mechanisms and fine localization of regulatory loci has achieved significant success, especially in highly complex genomic tree species [14,48,49]. Deciphering the haplotype genome of ginkgo will greatly help to resolve its important metabolic mechanisms, such as flavonoids, lactones, and polyprenols.

In summary, the difference between haploid and diploid ginkgo was mainly related to the reduced dosage of relevant genes. This work not only laid the foundation for the study of haploid ginkgo systems, it was also the first report to fully exploit haploid ginkgo genetic differences.

4. Materials and Methods

4.1. Plant Material and Sample Preparation

Haploid and diploid ginkgo plants were obtained from the Botanical Garden of Masaryk University. We grafted scions of 3 haploid ginkgo cultivars (‘Clica’, ‘Rocky’, and ‘Baldii’) and 3 diploid ginkgo cultivars (‘David’, ‘Tit’, and ‘Maronia’) on the same ginkgo cultivar in the Ginkgo Germplasm Resource Garden of Nanjing Forestry University (119.19 E, 31.58 N) in March 2020. All the scions were of the same length and had two buds. All scions were harvested from the first-year lateral branches of each maternal plant. The environmental conditions for the growth of the plants in the entire germplasm nursery were maintained in a group without significant adversity.

Ginkgo leaf samples for RNA isolation was collected after grafting the plants. The leaf samples were stored at $-80\text{ }^{\circ}\text{C}$ after flash freezing in liquid nitrogen.

4.2. RNA Sequencing and Data Analysis

Total RNA was extracted using the TRIzol reagent (Invitrogen Scientific, Inc., Carlsbad, CA, USA) in accordance with the manufacturer’s protocol. The quality of the total RNA of the sample was detected using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA) with an RNA integrity number (RIN) value of ≥ 7 required for all samples. After sequencing, the sequences were filtered to obtain high quality clean reads using fastp by removing the low quality bases and adapters in the sequences. Thus, data were further filtered using fastp [50] (version 0.18.0) to acquire high-quality clean reads, and the remaining clean reads were mapped to the most recent ginkgo reference genome [47] using HISAT2.4 [51]. The ribosomal RNA (rRNA) database was mapped using the short read alignment program Bowtie2 [52] (version 2.2.8). The mapped reads of each sample were assembled using StringTie [53,54] v1.3.1, using a reference-based approach. DESeq2 [55] software is used to analyze RNA differential expression between two separate groups (and edgeR [56] between two samples).

The gene sequencing method used for GSEA was Signal2Noise. FDR < 0.25 and normalized $p < 0.05$ were the filter conditions, and other parameter settings were kept as default values.

Bioinformatic analysis was performed using Omicsmart, a real-time interactive online platform for data analysis (<http://www.omicsmart.com>) (accessed on 2 May 2022).

The raw sequencing data from Illumina were uploaded to the NCBI Short Reads Archive database with the accession number SRP337737.

4.3. Determination of Flavonoid Content and Photosynthetic Rate of Ginkgo

We collected fully unfolded mature leaves and determined ginkgo flavonoid content using high-performance liquid chromatography (HPLC) [27]. First, the ground blade sample (0.5 g) was refluxed for 2 h with 50 mL of chloroform in a Soxhlet extractor and then evaporated to dryness. The dried residue was then extracted with 50 mL of methanol for 4 h, at 80 °C. After cooling to room temperature, the eluent was diluted to 50 mL with methanol, and the concentrations of quercetin, kaempferol, and Isorhamnetin were measured by HPLC. The following conditions were set for HPLC (Waters 1525, Canby, OR, USA): The mobile phase was a 1.0 mL/min solution of methanol and 0.4 percent H₃PO₄ (56:44, *v/v*); the column temperature was 30 °C; the sample was detected at 360 nm. The following conditions were set for HPLC (Waters 1525, USA): The mobile phase was a 1.0 mL/min solution of methanol and 0.4 percent H₃PO₄ (56:44, *v/v*). Total flavonoid content = (quercetin + kaempferol + isorhamnetin content) × 2.51 [57].

The photosynthetic rate of ginkgo was measured by CIRAS-3 (PP Systems, Amesbury, MA, USA) on a clear and windless morning in August 2021. The functional leaves of the same position were selected for each plant, and the determination was repeated three times. The photosynthetically effective radiation of the LED light source was 1200 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, consisting of 90% red light, 5% blue light and 5% white light.

4.4. Quantitative Real-Time PCR (qRT-PCR)

We used qRT-PCR to confirm the expression of 14 DEGs in the flavonoid metabolic and photosynthesis metabolic pathways. Detailed primer information is provided in Supplementary File S9. The $2^{-\Delta\Delta C_t}$ technique was used to calculate relative expression levels [58]. *GbCHS* (F: CAAGCGCATGTGCGACAAGT, R: CACCTCCACCACCACCATGT) was used as an endogenous control [59]. Each reaction in every experiment was repeated three times.

5. Conclusions

In this study, we found that haploid ginkgos exhibited lower growth potential, photosynthesis and flavonoid accumulation capacity through homogeneous garden trials of haploid and diploid ginkgos. The deepened transcriptome database revealed 2485 down-regulated DEGs expressed in haploid ginkgo, representing 53.3% of the whole DEGs and 8.9% of the whole genome. Although key genes in these metabolic pathways were expressed both up- and down-regulated in haploids, the gene dosage showed an overall decreasing trend by GSEA. Reduced gene dosage was the main factor leading to various differences in haploid ginkgos. This study provides a theoretical basis for the in-depth study and resource utilization of haploid ginkgo.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23168958/s1>.

Author Contributions: Methodology, Y.H.; software, Y.H.; formal analysis, Y.H.; data curation, G.L., B.W. and X.G.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H., P.Š. and Q.G.; project administration and funding acquisition, Q.G. 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 31971648 and the China Scholarship Council, grant number 202108320301.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Acknowledgments: We thank the Co-Innovation Center for Sustainable Forestry in Southern China for allowing us to use their research facilities.

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

References

- Wendel, J.F. Genome evolution in polyploids. *Plant Mol. Biol.* **2000**, *42*, 225–249. [[CrossRef](#)] [[PubMed](#)]
- Chung, M.Y.; Kim, C.Y.; Min, J.S.; Lee, D.-J.; Naing, A.H.; Chung, J.D.; Kim, C.K. In vitro induction of tetraploids in an interspecific hybrid of *Calanthe* (*Calanthe discolor* × *Calanthe sieboldii*) through colchicine and oryzalin treatments. *Plant Biotechnol. Rep.* **2014**, *8*, 251–257. [[CrossRef](#)]
- Li, Y.; Yang, J.; Song, L.; Qi, Q.; Du, K.; Han, Q.; Kang, X. Study of variation in the growth, photosynthesis, and content of secondary metabolites in *Eucommia* triploids. *Trees-Struct. Funct.* **2019**, *33*, 817–826. [[CrossRef](#)]
- Hu, Y.; Sun, D.; Hu, H.; Zuo, X.; Xia, T.; Xie, J. A comparative study on morphological and fruit quality traits of diploid and polyploid carambola (*Averrhoa carambola* L.) genotypes. *Sci. Hortic.* **2021**, *277*, 109843. [[CrossRef](#)]
- Zhang, W.; Hao, H.; Ma, L.; Zhao, C.; Yu, X. Tetraploid muskmelon alters morphological characteristics and improves fruit quality. *Sci. Hortic.* **2010**, *125*, 396–400. [[CrossRef](#)]
- Guo, H.; Mendrikahy, J.N.; Xie, L.; Deng, J.; Lu, Z.; Wu, J.; Li, X.; Shahid, M.Q.; Liu, X. Transcriptome analysis of neo-tetraploid rice reveals specific differential gene expressions associated with fertility and heterosis. *Sci. Rep.* **2017**, *7*, 40139. [[CrossRef](#)]
- Luo, G.; Xue, L.; Xu, W.; Zhao, J.; Wang, J.; Ding, Y.; Luan, K.; Lei, J. Breeding decaploid strawberry with improved cold resistance and fruit quality. *Sci. Hortic.* **2019**, *251*, 1–8. [[CrossRef](#)]
- Weber, D.F. Today's Use of Haploids in Corn Plant Breeding. In *Advances in Agronomy*; Spark, D.L., Ed.; Elsevier: Amsterdam, The Netherlands, 2014; Volume 123, pp. 123–144.
- Dwivedi, S.L.; Britt, A.B.; Tripathi, L.; Sharma, S.; Upadhyaya, H.D.; Rodomiro, O. Haploids: Constraints and opportunities in plant breeding. *Biotechnol. Adv. Int. Rev. J.* **2015**, *33*, 812–829. [[CrossRef](#)]
- Dunwell, J.M. Haploids in flowering plants: Origins and exploitation. *Plant Biotechnol. J.* **2010**, *8*, 377–424. [[CrossRef](#)] [[PubMed](#)]
- Isakov, Y.N.; Butorina, A.K.; Muraya, L.S. Discovery of spontaneous haploids in *Pinus silvestris* and the prospects of their using in forest genetics and selection. *Genetika* **1981**, *17*, 701–707.
- Blakeslee, A.F.; Belling, J.; Farnham, M.E.; Bergner, A.D. A Haploid Mutant in the Jimson Weed, “*Datura Stramonium*”. *Science* **1922**, *55*, 646–647. [[CrossRef](#)] [[PubMed](#)]
- Von Dassow, P.; Ogata, H.; Probert, I.; Wincker, P.; Da Silva, C.; Audic, S.; Claverie, J.M.; de Vargas, C. Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. *Genome Biol.* **2009**, *10*, R114. [[CrossRef](#)] [[PubMed](#)]
- Li, Y.; Wei, H.; Yang, J.; Du, K.; Li, J.; Zhang, Y.; Qiu, T.; Liu, Z.; Ren, Y.; Song, L.; et al. High-quality de novo assembly of the *Eucommia ulmoides* haploid genome provides new insights into evolution and rubber biosynthesis. *Hortic. Res.* **2020**, *7*, 183. [[CrossRef](#)]
- Zhou, Q.; Tang, D.; Huang, W.; Yang, Z.; Zhang, Y.; Hamilton, J.P.; Visser, R.G.F.; Bachem, C.W.B.; Robin Buell, C.; Zhang, Z.; et al. Haplotype-resolved genome analyses of a heterozygous diploid potato. *Nat. Genet.* **2020**, *52*, 1018–1023. [[CrossRef](#)]
- Kelliher, T.; Starr, D.; Richbourg, L.; Chintamanani, S.; Delzer, B.; Nuccio, M.L.; Green, J.; Chen, Z.; McCuiston, J.; Wang, W.; et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* **2017**, *542*, 105–109. [[CrossRef](#)]
- Wang, N.; Xia, X.; Jiang, T.; Li, L.; Zhang, P.; Niu, L.; Cheng, H.; Wang, K.; Lin, H. In planta haploid induction by genome editing of DMP in the model legume *Medicago truncatula*. *Plant Biotechnol. J.* **2022**, *20*, 22–24. [[CrossRef](#)]
- Zhong, Y.; Chen, B.; Wang, D.; Zhu, X.; Li, M.; Zhang, J.; Chen, M.; Wang, M.; Riksen, T.; Liu, J.; et al. In vivo maternal haploid induction in tomato. *Plant Biotechnol. J.* **2022**, *20*, 250–252. [[CrossRef](#)]
- Sun, Y.H.; Wang, Y.R.; Li, Y.Y.; Jiang, J.Z.; Yang, N.; Niu, C.; Li, Y. Effects of colchicine treatment on the microtubule cytoskeleton and total protein during microsporogenesis in *Ginkgo biloba* L. *Pak. J. Bot.* **2015**, *47*, 159–170.
- Zhao, Y.; Paule, J.; Fu, C.; Koch, M.A. Out of China: Distribution history of *Ginkgo biloba* L. *Taxon* **2010**, *59*, 495–504. [[CrossRef](#)]
- McElwain, J.C. Ginkgo The Tree That Time Forgot. *Science* **2013**, *340*, 812–813. [[CrossRef](#)]
- Efferth, T.; Koch, E. Complex Interactions between Phytochemicals. The Multi-Target Therapeutic Concept of Phytotherapy. *Curr. Drug Targets* **2011**, *12*, 122–132. [[CrossRef](#)]
- Zhao, J.; Yang, J.; Xie, Y. Improvement strategies for the oral bioavailability of poorly water-soluble flavonoids: An overview. *Int. J. Pharm.* **2019**, *570*, 118642. [[CrossRef](#)]
- Wu, D.; Feng, J.; Lai, M.; Ouyang, D.; Yu, W.; Wang, G.; Cao, F.; Jacobs, D.F.; Zeng, S. Combined application of bud and leaf growth fertilizer improves leaf flavonoids yield of *Ginkgo biloba*. *Ind. Crops Prod.* **2020**, *150*, 112379. [[CrossRef](#)]
- Jun, N.; Juan, H.; Zhifang, J.; Xiaori, Z.; Lixiang, D.; Xiuli, Y.; Zhehang, S.; Wenya, X.; Zhikun, W.; Maojun, X. NaCl Induces Flavonoid Biosynthesis through a Putative Novel Pathway in Post-harvest Ginkgo Leaves. *Front. Plant Sci.* **2017**, *8*, 920.
- Wang, G.; Cao, F.; Chang, L.; Guo, X.; Wang, J. Temperature has more effects than soil moisture on biosynthesis of flavonoids in *Ginkgo* (*Ginkgo biloba* L.) leaves. *New For.* **2014**, *45*, 797–812. [[CrossRef](#)]
- Xu, Y.; Wang, G.; Cao, F.; Zhu, C.; Wang, G.; El-Kassaby, Y.A. Light intensity affects the growth and flavonol biosynthesis of *Ginkgo* (*Ginkgo biloba* L.). *New For.* **2014**, *45*, 765–776. [[CrossRef](#)]

28. Ying, G.; Cga, B.; Mwa, B.; Fffa, B.; Ek, C.; Tw, C.; Gwa, B. Metabolome and transcriptome analyses reveal flavonoids biosynthesis differences in *Ginkgo biloba* associated with environmental conditions—ScienceDirect. *Ind. Crops Prod.* **2020**, *158*, 112963.
29. Ni, J.; Dong, L.; Jiang, Z.; Yang, X.; Chen, Z.; Wu, Y.; Xu, M. Comprehensive transcriptome analysis and flavonoid profiling of Ginkgo leaves reveals flavonoid content alterations in day–night cycles. *PLoS ONE* **2018**, *13*, e0193897. [[CrossRef](#)]
30. Wu, Y.; Guo, J.; Qi, Z.; Xin, Y.; Wang, G.; Xu, L.A. De novo transcriptome analysis revealed genes involved in flavonoid biosynthesis, transport and regulation in *Ginkgo biloba*. *Ind. Crops Prod.* **2018**, *124*, 226–235. [[CrossRef](#)]
31. Smarda, P.; Vesely, P.; Smerda, J.; Bures, P.; Knappek, O.; Chytra, M. Polyploidy in a “living fossil” *Ginkgo biloba*. *New Phytol.* **2016**, *212*, 11–14. [[CrossRef](#)]
32. Smarda, P.; Horova, L.; Knappek, O.; Dieck, H.; Dieck, M.; Razna, K.; Hrubik, P.; Orloci, L.; Papp, L.; Vesela, K.; et al. Multiple haploids, triploids, and tetraploids found in modern-day “living fossil” *Ginkgo biloba*. *Hortic. Res.* **2018**, *5*, 55. [[CrossRef](#)]
33. Lin, Z.; Pu, H.; Duan, Z.; Li, Y.; Liu, B.; Zhang, Q.; Li, W.; Jean-David, R.; Liu, L.; Peng, L. Nucleus-Encoded Protein BFA1 Promotes Efficient Assembly of the Chloroplast ATP Synthase Coupling Factor 1. *Plant Cell* **2018**, *30*, 1770–1788.
34. Tamayo, P.; Steinhardt, G.; Liberzon, A.; Mesirov, J.P. The limitations of simple gene set enrichment analysis assuming gene independence. *Stat. Methods Med. Res.* **2016**, *25*, 472–487. [[CrossRef](#)]
35. Tilford, C.A.; Siemers, N.O. Gene Set Enrichment Analysis. In *Protein Networks and Pathway Analysis*; Nikolsky, Y., Bryant, J., Eds.; Methods in Molecular Biology: New York, NY, USA, 2009; Volume 563, pp. 99–121.
36. Chen, T.T.; Sheng, Y.; Hao, Z.D.; Long, X.F.; Fu, F.F.; Liu, Y.; Tang, Z.H.; Ali, A.; Peng, Y.; Lu, L.; et al. Transcriptome and proteome analysis suggest enhanced photosynthesis in tetraploid *Liriodendron sino-americanum*. *Tree Physiol.* **2021**, *41*, 1953–1971. [[CrossRef](#)]
37. Tokumoto, Y.; Kajiura, H.; Takeno, S.; Harada, Y.; Suzuki, N.; Hosaka, T.; Gyokusen, K.; Nakazawa, Y. Induction of tetraploid hardy rubber tree, *Eucommia ulmoides*, and phenotypic differences from diploid. *Plant Biotechnol.* **2016**, *33*, 51–57. [[CrossRef](#)]
38. Wu, Y.; Sun, Y.; Sun, S.; Li, G.; Wang, J.; Wang, B.; Lin, X.Y.; Huang, M.; Gong, Z.Y.; Sanguinet, K.A.; et al. Aneuploidization under segmental allotetraploidy in rice and its phenotypic manifestation. *Theor. Appl. Genet.* **2018**, *131*, 1273–1285. [[CrossRef](#)]
39. Mo, L.; Chen, J.H.; Chen, F.; Xu, Q.W.; Tong, Z.K.; Huang, H.H.; Dong, R.H.; Lou, X.Z.; Lin, E.P. Induction and characterization of polyploids from seeds of *Rhododendron fortunei* Lindl. *J. Integr. Agric.* **2020**, *19*, 2016–2026. [[CrossRef](#)]
40. Shmeit, Y.H.; Fernandez, E.; Novy, P.; Kloucek, P.; Orosz, M.; Kokoska, L. Autopolyploidy effect on morphological variation and essential oil content in *Thymus vulgaris* L. *Sci. Hortic.* **2020**, *263*. [[CrossRef](#)]
41. Ma, Y.; Xue, H.; Zhang, L.; Zhang, F.; Ou, C.Q.; Wang, F.; Zhang, Z.H. Involvement of Auxin and Brassinosteroid in Dwarfism of Autotetraploid Apple (*Malus x domestica*). *Sci. Rep.* **2016**, *6*, 26719. [[CrossRef](#)]
42. Allario, T.; Brumos, J.; Colmenero-Flores, J.M.; Tadeo, F.; Froelicher, Y.; Talon, M.; Navarro, L.; Ollitrault, P.; Morillon, R. Large changes in anatomy and physiology between diploid Rangpur lime (*Citrus limonia*) and its autotetraploid are not associated with large changes in leaf gene expression. *J. Exp. Bot.* **2011**, *62*, 2507–2519. [[CrossRef](#)]
43. Adams, K.L. Evolution of duplicate gene expression in polyploid and hybrid plants. *J. Hered.* **2007**, *98*, 136–141. [[CrossRef](#)] [[PubMed](#)]
44. Udall, J.A.; Wendel, J.F. Polyploidy and crop improvement. *Crop Sci.* **2006**, *46*, S3–S14. [[CrossRef](#)]
45. Kashkush, K.; Feldman, M.; Levy, A.A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* **2003**, *33*, 102–106. [[CrossRef](#)] [[PubMed](#)]
46. Nakagami, H.; Pitzschke, A.; Hirt, H. Emerging MAP kinase pathways in plant stress signalling. *Trends Plant Sci.* **2005**, *10*, 339–346. [[CrossRef](#)] [[PubMed](#)]
47. Liu, H.; Wang, X.; Wang, G.; Cui, P.; Wu, S.; Ai, C.; Hu, N.; Li, A.; He, B.; Shao, X.; et al. The nearly complete genome of *Ginkgo biloba* illuminates gymnosperm evolution. *Nat. Plants* **2021**, *7*, 748–756. [[CrossRef](#)] [[PubMed](#)]
48. Hasing, T.; Tang, H.; Brym, M.; Khazi, F.; Huang, T.; Chambers, A.H. A phased *Vanilla planifolia* genome enables genetic improvement of flavour and production. *Nat. Food* **2020**, *1*, 811–819. [[CrossRef](#)]
49. Tian, Y.; Thrimawithana, A.; Ding, T.; Guo, J.; Gleave, A.; Chagne, D.; Ampomah-Dwamena, C.; Ireland, H.S.; Schaffer, R.J.; Luo, Z.; et al. Transposon insertions regulate genome-wide allele-specific expression and underpin flower colour variations in apple (*Malus* spp.). *Plant Biotechnol. J.* **2022**, *20*, 1285–1297. [[CrossRef](#)]
50. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, 884–890. [[CrossRef](#)]
51. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [[CrossRef](#)]
52. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
53. Perteau, M.; Perteau, G.M.; Antonescu, C.M.; Chang, T.-C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [[CrossRef](#)] [[PubMed](#)]
54. Perteau, M.; Kim, D.; Perteau, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650–1667. [[CrossRef](#)] [[PubMed](#)]
55. 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](#)] [[PubMed](#)]
56. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)]
57. Chinese Pharmacopoeia Commission. 2015. Available online: <http://www.chp.org.cn> (accessed on 15 September 2021).

-
58. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2013**, *25*, 402–408. [[CrossRef](#)]
 59. Zhou, T.; Yang, X.; Fu, F.; Wang, G.; Cao, F. selection of suitable reference genes based on transcriptomic data in ginkgo biloba under different experimental conditions. *Forests* **2020**, *11*, 1217. [[CrossRef](#)]