Article

Genome-Wide Analysis of Alternative Splicing and Non-Coding RNAs Reveal Complicated Transcriptional Regulation in *Cannabis sativa* L.

Bin Wu [1],[†], Yanni Li [1],[†], Jishuang Li [1], Zhenzhen Xie [1], Mingbao Luan [2], Chunsheng Gao [2], Yuhua Shi [3] and Shilin Chen [3],*

[1] Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China; bwu@implad.ac.cn (B.W.); liyanni_hua1126@126.com (Y.L.); ljs1378905466@126.com (J.L.); zhzhxie126@126.com (Z.X.)
[2] Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha 410205, China; luanmingbao2002@126.com (M.L.); csgao06@163.com (C.G.)
[3] Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China; yhshi@icmm.ac.cn
* Correspondence: slchen@icmm.ac.cn
† These authors contributed equally to this work.

Abstract: It is of significance to mine the structural genes related to the biosynthetic pathway of fatty acid (FA) and cellulose as well as explore the regulatory mechanism of alternative splicing (AS), microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) in the biosynthesis of cannabinoids, FA and cellulose, which would enhance the knowledge of gene expression and regulation at post-transcriptional level in *Cannabis sativa* L. In this study, transcriptome, small RNA and degradome libraries of hemp 'Yunma No.1' were established, and comprehensive analysis was performed. As a result, a total of 154, 32 and 331 transcripts encoding key enzymes involved in the biosynthesis of cannabinoids, FA and cellulose were predicted, respectively, among which AS occurred in 368 transcripts. Moreover, 183 conserved miRNAs, 380 *C. sativa*-specific miRNAs and 7783 lncRNAs were predicted. Among them, 70 miRNAs and 17 lncRNAs potentially targeted 13 and 17 transcripts, respectively, encoding key enzymes or transporters involved in the biosynthesis of cannabinoids, cellulose or FA. Finally, the crosstalk between AS and miRNAs or lncRNAs involved in cannabinoids and cellulose was also predicted. In summary, all these results provided insights into the complicated network of gene expression and regulation in *C. sativa*.

Keywords: *Cannabis sativa* L.; alternative splicing; non-coding RNAs; gene expression and regulation

1. Introduction

Recently, *Cannabis sativa* L. has attracted worldwide attention due to its wide use in medicine, food, health products, textiles and other fields [1,2]. The main products of *C. sativa* are derived from cannabinoids, fatty acid (FA) and cellulose [3–5]. Cannabinoids are types of *C. sativa*-specific secondary metabolites rich in female flowers [6,7]. To date, more than 100 cannabinoid compounds have been found in *C. sativa*, of which Δ9-tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA) are the major components [8]. Both THCA and CBDA have functions in relieving pain, keeping calm, preventing vomiting, regulating immunity and reducing artery occlusion [9,10]. However, THCA has hallucinogenic and addictive effects [11,12]. Content and composition of cannabinoids are highly variable among *C. sativa* plants. According to the content and ratio of THCA and CBDA, *C. sativa* can be mainly divided into marijuana and hemp. Marijuana contains high THCA and low CBDA, while hemp is the opposite. Besides CBDA, hemp is cultivated for fiber and oil. Hemp seed is a kind of nutritious food, rich in FA [13]. Moreover, the proportion of ω-6 and ω-3 FAs is 3:1, consistent with the recommended

balance for human health [14]. The cultivation area and the export of products of hemp in China account for about 50% and 25% of the world, respectively [15].

With the release of the draft genome of *C. sativa*, the biosynthetic pathway of cannabinoids has been deciphered and is involved in the pathways of hexanoate, plastid-localized 2-C-methyl-d-erythritol 4-phosphate (MEP) and geranyl pyrophosphate (GPP) [16–18]. The biosynthetic pathway of FA has been studied in some plants [19,20], but there is no report on *C. sativa* so far. The biosynthesis of cellulose is very complex in plants. Besides cellulose synthase (CESA), some non-CESA proteins are also involved in the process [21]. However, the genes related to the biosynthesis of cellulose in *C. sativa* have not been mined yet. In brief, the regulation in the biosynthesis of cannabinoids, FA and cellulose in *C. sativa* is still unknown to date.

It is known that gene expression and regulation consist of many layers. Alternative splicing (AS) is a significant mechanism for regulating the expression of genes in eukaryotes [22]. AS refers to one mRNA precursor producing diverse transcripts, which is thought of as an important post-transcriptional mechanism for increasing the diversity of transcripts and proteins [23]. AS plays a major role in affecting protein function; therefore, it is necessary to systematically identify AS events. However, AS has not been reported in *C. sativa.*

Non-coding RNAs, especially for microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), have been comprehensively studied using bioinformatic and experimental approaches in plants [24,25]. However, they are not reported in *C. sativa*. MiRNAs are types of 20–24 nucleotide non-coding RNAs, usually originating from single-stranded transcripts with imperfect stem-loop structure [26], which cleavage or inhibit translation of target genes at posttranscriptional level or through methylation modification at the transcriptional level [27–29]. MiRNAs function as gene regulators during the plant growth process [30–32] and are also involved in adaptive responses to abiotic stresses [33,34] and synthesis of secondary metabolites [35–37]. LncRNAs have become a popular research topic in molecular biology in recent years. LncRNAs display a key regulatory role in plant cell development, which can be trans-acting for target genes in a sequence complementary manner [38] or cis-acting in other mechanisms [39].

It was reported that AS and miRNAs could crosstalk, which increased the complexity of gene expression and regulation, emerging as one of the hot topics in molecular biology. On the one hand, the AS isoform affected the biogenesis of miRNA and lncRNA. For example, there is a kind of tissue-specific AS regulating the biogenesis of miR-412 involved in cell death [40]. MiR846 and miR842 are regulated by abscisic acid, originating from different AS isoforms [41]. On the other hand, miRNA targeted partial AS isoforms, decreasing the regulation by miRNA [42].

In this study, the structural and transporter-coding genes related to the biosynthetic pathway of FA and cellulose were mined in *C. sativa* for the first time. Then, AS events, miRNAs and lncRNAs were analyzed, and the potential regulation in the biosynthesis of cannabinoids, FAs and cellulose was predicted. Moreover, the crosstalk between AS and miRNAs or lncRNAs was revealed. Taken together, all these results will offer candidate genes related to the biosynthetic pathway of cannabinoids, FA and cellulose and enhance the knowledge of gene expression and regulation in hemp.

2. Results

2.1. RNA-Seq Library Construction and Analysis

'Yunma No.1' is the first industrial hemp cultivar bred by the Yunnan Academy of Agricultural Sciences. Its THCA content is <0.3%, complying with the internationally adopted EU standards for hemp [43]. To obtain transcriptomic information of 'Yunma No.1', a strand-specific RNA-Seq library was constructed and sequenced by 90 bp paired-end (PE) sequencing. After filtering the low-quality reads, in total, ~66 million clean reads were obtained. More than 83% of them could be mapped to the published *C. sativa* 'Purple Kush' genome using tophat with the default parameter [44]. Then, the 735,445,652 mapped reads were used for assembly by Cufflinks software [45]. To predict the gene functions,

the 67,035 assembly transcripts were annotated by NCBI nonredundant protein sequences (NR) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases [46,47].

2.2. Structural and Transporter Genes Related to the Biosynthetic Pathway of Cannabinoids, FA and Cellulose

The genes related to cannabinoids, FA and cellulose metabolism were focused on because the main products of *C. sativa* are derived from cannabinoids, FA and cellulose [3–5].

Based on the NR annotation, a total of 154 transcripts involved in the biosynthesis of cannabinoids, via the hexanoate, MEP and GPP pathways, were obtained. In the hexanoate pathway, the Delta12-oleic acid desaturase gene was identified [48], which contains one transcript; the Lipoxygenase (LOX) gene contains eight transcripts, and the Acyl-activating enzyme (AAE) gene contains 87 transcripts. In MEP pathway, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS), 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (HDS) and 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR) genes were identified, which contain 6, 13, 7, 2, 3, 15 and 3 transcripts, respectively. In the GPP pathway, the geranyl diphosphate synthase (GPPS) gene was identified, which contains one transcript. In the cannabinoid biosynthetic pathway, the AAE gene catalyzed the hexanoate to form hexanoyl-CoA, which provided the olivetol synthase (OLS) forming olivetolic acid with substrate. The OLS gene was identified, containing one transcript. The MEP pathway supplies substrates for GPPS, and the prenyl side-chain originates from the prenyltransferase (PT) gene to form cannabigerolic acid (CBGA) [49]. GPPS and PT genes contained one and eight transcripts, respectively. Finally, THCA and CBDA are separately synthesized through THCA synthase (THCAS) and CBDA synthase (CBDAS) (Table S1, Figure 1).

In the FA biosynthetic pathway, firstly, the pyruvate oxidative was decarboxylated by the pyruvate dehydrogenase complex (PDHC) to form acetyl-CoA in the plastid, which was reversely catalyzed by acetyl-CoA carboxylase (ACCase) to produce malonyl-CoA [19]. In this study, 18 transcripts of PAHC and eight transcripts of ACCase were identified. For the malonyl-CoA, ACP transacylase (MCAAT) transferred malonyl-CoA to malonyl-acyl carrier protein (ACP), which gradually converted into saturated FAs by β-ketoacyl-ACP synthase (KAS), β-ketoacyl-ACP reductase (KAR), hydroxyacyl-ACP dehydrase (HAD) and enoyl-ACP reductase (EAR). The KAR gene was identified as containing one transcript. The butyryl-ACP and malonyl-ACP were catalyzed by KAS I and KAS II to form 16-carbon palmitoyl-ACP (C16:0-ACP) and 18-carbon stearoyl-ACP (C18:0-ACP). C18:0-ACP was efficiently catalyzed by stearoyl-ACP desaturase (SAD) to produce C18:1-ACP. Next, oleate desaturase (chloroplast-type) (FAD6) catalyzed C18:1-ACP to form C18:2-ACP, which was finally catalyzed by linoleate desaturase (chloroplast-type) (FAD7/8) to form C18:3-ACP [50,51]. Both FAD6 and FAD7/8 genes contained one transcript. Then, C16:0-ACP, C18:0-ACP, C18:1-ACP, C18:2-ACP and C18:3-ACP could be further catalyzed by acyl-ACP thioesterase A (FATA), acyl-ACP thioesterase B (FATB) and palmitoyl-CoA hydrolase (PCH), respectively [52], to produce free FAs (including C16:0, C18:0, C18:1, C18:2 and C18:3) [20,53]. Three unigenes encoding the FATB enzyme were identified (Table S2, Figure 2). Among these free FAs, oleic acid (ω-9 C18:1) and linoleic acid (ω-6 C18:2) as precursor substances are involved in the hexanoate pathway of the cannabinoid synthesis pathway to synthesize the hexanoate in hemp [48,54–56].
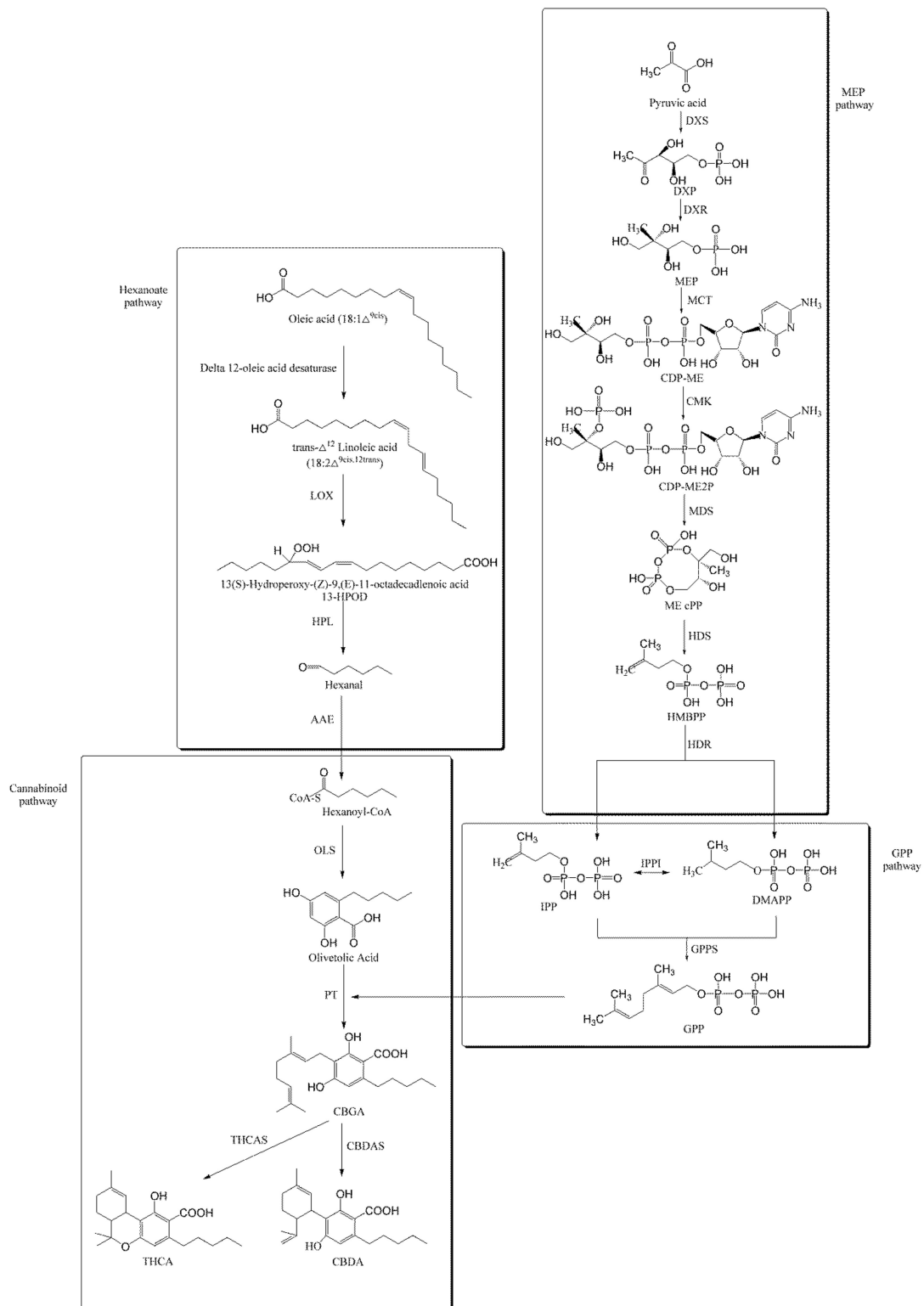
Figure 1. Proposed THCA and CBDA biosynthetic pathway in *C. sativa*. LOX: lipoxygenase; HPL: hydroperoxide lyase; AAE: acyl-activating enzyme; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; DXR: 1-deoxy-D-xylulose 5-phosphate reductoisomerase; MCT: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MDS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; IPPI: Isopentenyl diphosphate isomerase; GPPS: Geranyl diphosphate synthase; OLS: olivetol synthase; PT: prenyltransferase; THCAS: tetrahydrocannabinolic acid synthase; CBDAS: cannabidiolic acid synthase.
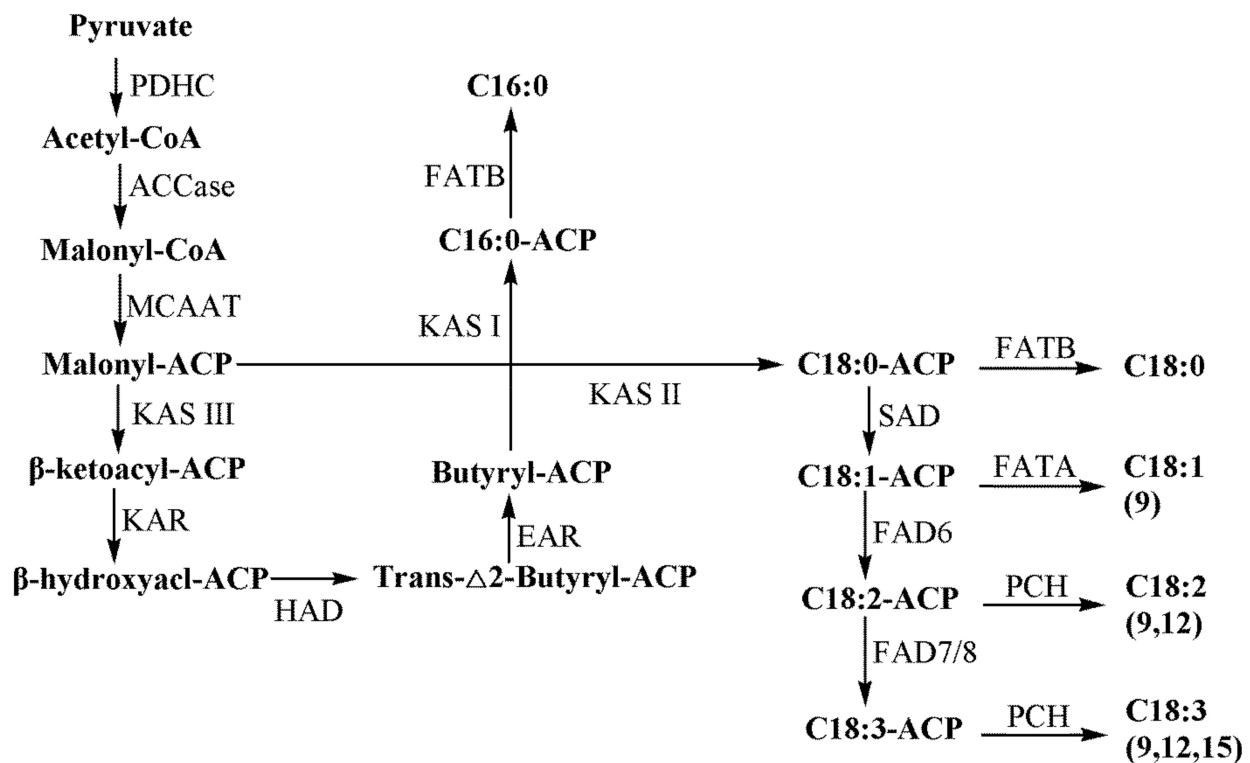
Figure 2. Proposed fatty acid biosynthetic pathway in *C. sativa* [20]. PDHC: pyruvate dehydrogenase complex; ACCase: acetyl-CoA carboxylase; MCAAT: malonyl-CoA: ACP transacylase; KAS: β-ketoacyl-ACP synthase; KAR: β-ketoacyl-ACP reductase; HAD: hydroxyacyl-ACP dehydrase; EAR: enoyl-ACP reductase; SAD: stearoyl-ACP desaturase; FAD6: oleate desaturase (chloroplast-type); FAD7/8: linoleate desaturase (chloroplast-type); FATA/B: acyl-ACP thioesterase A/B; PCH: palmitoyl-CoA hydrolase.

In addition, 331 transcripts potentially involved in the cellulose metabolism were also identified. Cellulose synthase (CESA) can synthesize cellulose at the plasma membrane in land plants [57], which contain 27 transcripts. Sucrose synthase (SuSy) catalyzes UDP into UDP-glucose for each CESA can synthesize a single glucan chain [58], which contains 18 transcripts. The UDP-galactose and UDP-glucose transporter are reported that make substrates available for polysaccharide biosynthesis and have an important impact on maintaining cell wall integrity [59,60]. There were seven and 28 transcripts classified into UDP-galactose and UDP-glucose transporters, respectively. In the process of secretory vesicles targeting the SCW domain, the motor proteins kinesin-4 and kinesin-like proteins are involved in cell transport in microtubules [61,62], which contained four and 175 transcripts. There was a number of non-CESA proteins also involved in cellulose production, including COBRA-like protein (COBL), endo-1,4-beta-glucanase, chitinase-like protein (CTL), alpha-glucosidase I, VILLIN, katanin-like protein, glycosyltransferase-like protein KOBITO1 and inositol polyphosphate 5-phosphatase (FRA) [63–72], which contained eight, one, four, one, two, three, one and three transcripts, respectively (Table S3).

2.3. Genome-Wide Identification of AS

AS events were identified using ASTALAVISTA with default options [73]. Results showed that 27,907 AS events were predicted, including 9773 intron retention (IntronR), 3404 exon skipping (ExonS), 15,970 alternative 3′ last exon (AltA) and 6388 alternative 5′ first exon (AltD) (Table S4) [74]. These results were consistent with the four AS types that account for the majority in plants [75]. To validate the authenticity of the predicted AS, five AS events were verified by RT-PCR. The electrophoresis results were in concordance with the RNA-Seq data (Figure S1).

Then, the AS events that occurred in structural genes related to cannabinoids, FA and cellulose were analyzed. Results showed that AS occurred in the cannabinoid biosynthetic structural genes *LOX*, *AAE*, *PT*, *DXR*, *MCT* and *HDR* (Table S5). AS also occurred in the FA biosynthetic structural genes, including *PDHC* and *ACCase* occurred AS (Table S6) as well as the cellulose biosynthetic structural genes, including *CESA*, *CSL*, UDP-galactose transporter, UDP-glucose transporter, kinesin-like protein, kinesin-4, COBRA-like protein, *CTL*, katanin-like protein, *FRA*, *SuSy* and *MAPKKK* (Table S7).

2.4. Identification of Regulatory Proteins in C. sativa

Regulatory proteins play a key role in various biological processes, such as vegetation growth, stress response and the synthesis of bioactive compounds [76,77]. To predict the regulatory proteins in *C. sativa*, the 48,799 assembly transcripts were compared with the sequences in the iTAK database. In total, 2012 tfs were identified including 63 families in which myb- > myb-related (152), bhlh (128), myb- > myb (118), c3h (114) and garp- > garp-g2-like (99) were the top five families. A total of 900 trs were found including 24 families, in which set (152), snf2 (128), others (127), phd (79) and gnat (74) were the top five families. A total of 2161 pks were identified, which included 114 families. Among them, the largest family is rlk-pelle_dlsv with 354 members (Table S8). Among them, some regulatory proteins were potentially involved in the cellulose metabolism. For example, the mitogen-activated protein kinase kinase (MAPKKK) could promote cell division, differentiation and the synthesis of cellulose in *Arabidopsis* [78], which contains 48 transcripts. Myb domain protein 46 (myb46) enhanced the content of cellulose by targeting several secondary wall nac regulators and *cesas* in *Arabidopsis* [79], which contained one transcript (Table S3).

2.5. Identification of miRNAs in C. sativa

The small RNA library was constructed using pooled RNA of stems, roots and leaves of *C. sativa* and sequenced by the HiSeq 2000 platform. After filtering, a total of 22,164,487 clean reads ranging from 18 to 30 nucleotide (nt) in size were obtained. Among them, the clean reads with counts ≥5 were further analyzed by the psRobot and Mireap software to predict the miRNAs in *C. sativa*. After checking manually, a total of 563 miRNAs were predicted. Among them, 183 miRNAs were identified as conserved miRNAs by comparing with miRbase, which could be divided into 36 families in which miR156 was the largest family, with 16 members. The remaining 380 miRNAs were adjudged to be *C. sativa*-specific miRNAs, which were grouped into 60 families. Among them, CsmiRNA-n4 was the largest family that possessed 34 members (Table S9).

2.6. Target Prediction of miRNAs in C. sativa

The psRNATarget software was used to predict the target genes of the miRNAs in *C. sativa*. As a result, the 178 conserved and 365 *C. sativa*-specific miRNAs were found to potentially target 1822 and 3951 transcripts, respectively (Table S10). These transcripts were involved in multiple functions. Among these target transcripts, 2673 of them were annotated by the KEGG database. KEGG enrichment showed that 15, 25, 160, 271 and 29 transcripts were involved in cellular processes, environmental information processing, genetic information processing, metabolism and organismal systems, respectively, in which the top five enrichment pathways were RecQ-mediated genome instability protein 1, CCR4-NOT transcription complex subunit 4, KAS II, GPI ethanolamine phosphate transferase 3 subunit O and alpha-mannosidase (Figure 3).

Then, we considered the target genes potentially involved in the biosynthesis of cannabinoids, FA and cellulose. As a result, a total of 21 miRNAs including four conserved families of miR156/miR477/miR5658 and 17 *C. sativa*-specific miRNAs including families of CsmiRNA-n28/n33/n34/n39/n57 were found to potentially target *AAE, DXR, DXS, HDR* and *LOX* structural genes in cannabinoid biosynthesis. MiR530 and CsmiRNA-n52 f targeted *ACCase* structural genes in FA biosynthesis. Nine conserved miRNAs families including miR1508, miR171, miR172, miR482 and miR5569 and 35 *C. sativa*-specific miRNAs

including families of CsmiRNA-n1/n4/n15/n22/n23/n24/n25/n27/n32/n34/n37/n56 were found to potentially target *CESA*, COBRA-like protein, *CSL*, *FRA*, kinesin-like protein, UDP-galactose transporter and *VILLIN* genes in the biosynthesis of cellulose (Table 1). The regulatory protein genes targeted by miRNAs were also investigated. Results showed that 218, 99 and 186 miRNAs potentially targeted 337 TFs, 148 TRs and 290 PKs, respectively (Table S11).
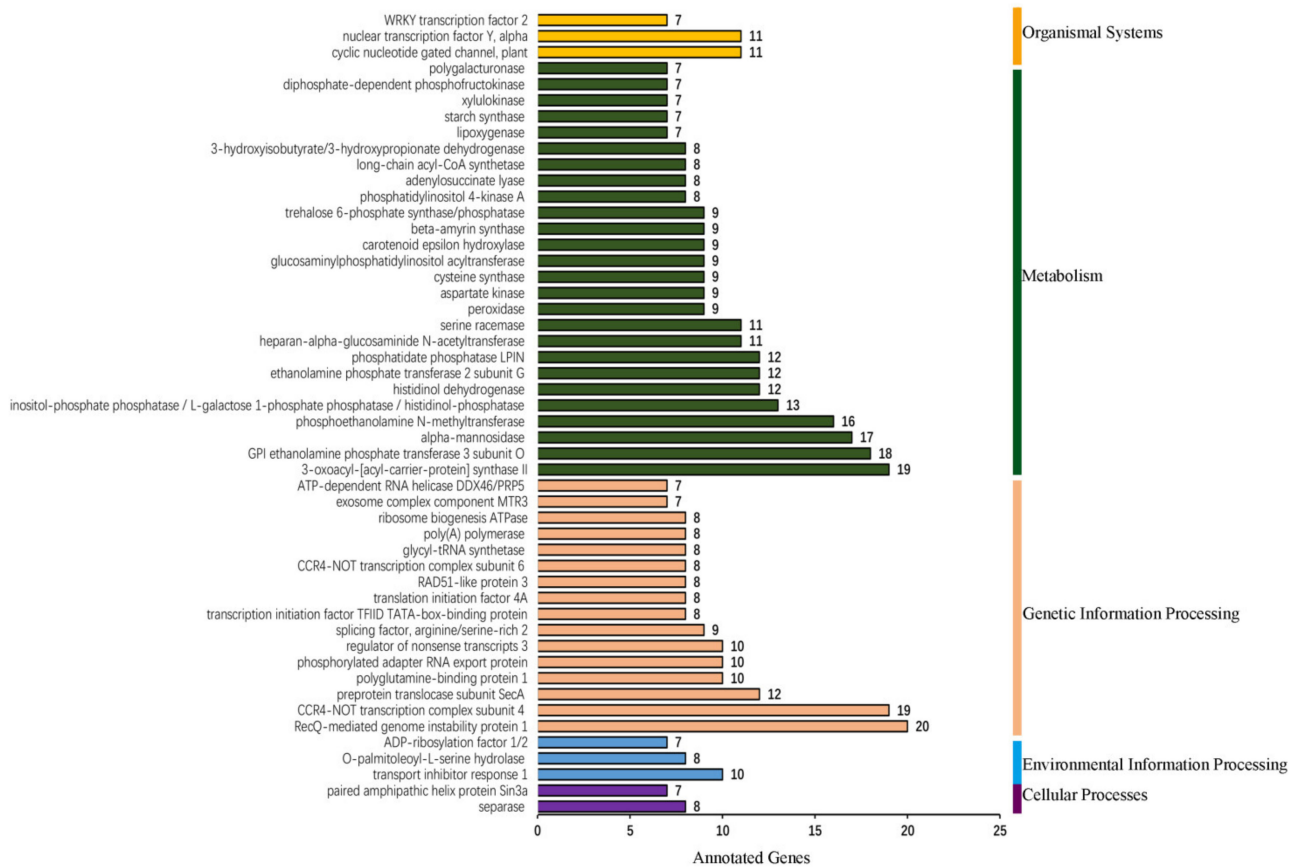
Figure 3. KEGG enrichment of miRNA targets.

Table 1. MiRNAs target structural genes potentially involved in cannabinoids, fatty acid and cellulose biosynthesis.

miRNA Name	Target ID	Target Annotation	Expectation	Pathway
CsmiRNA-n33j.1-3p	T_00033798		3	
	T_00033799		3	
	T_00033800		3	
	T_00007721		3	
	T_00007722		3	
	T_00069673		3	
	T_00069674		3	
	T_00069676		3	
miR5658	T_00030776	AAE	3	cannabinoids
	T_00030777		3	
	T_00030778		3	
	T_00030779		3	
miR477a-5p	T_00014806		3	
miR477b-5p			3	
miR477a-5p	T_00014807		3	
miR477b-5p			3	
miR156e-3p	T_00050643		1.5	
CsmiRNA-n39d.1-3p	T_00043759	DXR	3	
CsmiRNA-n39e.1			3	

Table 1. *Cont.*

miRNA Name	Target ID	Target Annotation	Expectation	Pathway
CsmiRNA-n34g.2	T_00038493		2.5	
CsmiRNA-n34a-5p			0	
CsmiRNA-n34b-5p			0.5	
CsmiRNA-n34c-5p			1	
CsmiRNA-n34d.1-5p			2	
CsmiRNA-n34d.2		DXS	3	
CsmiRNA-n34e.2-3p	T_00064406		0.5	
CsmiRNA-n34e.2-5p			0.5	
CsmiRNA-n34f.1-5p			0.5	
CsmiRNA-n34f.2-5p			3	
CsmiRNA-n34h.2-3p			1.5	
CsmiRNA-n57-5p			2.5	
CsmiRNA-n28	T_00018927		1.5	
	T_00018928	HDR	1.5	
	T_00018929		1.5	
miR156e-3p	T_00090667		3	
CsmiRNA-n33e-3p	T_00023464	LOX	2.5	
	T_00023465		2.5	
	T_00023466		2.5	
CsmiRNA-n52a			3	
CsmiRNA-n52d	T_00067318		3	
CsmiRNA-n52e			3	
CsmiRNA-n52f			3	
CsmiRNA-n52a			3	
CsmiRNA-n52d	T_00067319	ACCase	3	fatty acid
CsmiRNA-n52e			3	
CsmiRNA-n52f			3	
CsmiRNA-n52a			3	
CsmiRNA-n52d	T_00067320		3	
CsmiRNA-n52e			3	
CsmiRNA-n52f			3	
miR530-3P	T_00087904		3	
CsmiRNA-n27a.2			1	
CsmiRNA-n27c.2			2.5	
CsmiRNA-n27d-3p			2	
CsmiRNA-n27e-3p			2	
CsmiRNA-n27f	T_00016027		1.5	
CsmiRNA-n27g-3p			0.5	
CsmiRNA-n27i-3p			1.5	
CsmiRNA-n27l			0.5	
CsmiRNA-n27n-3p			0.5	
CsmiRNA-n27a.1			2.5	
CsmiRNA-n27b		COBL	0	cellulose
CsmiRNA-n27c.1			0.5	
CsmiRNA-n27d-5p			1	
CsmiRNA-n27e-5p			0	
CsmiRNA-n27g-5p			2.5	
CsmiRNA-n27h			3	
CsmiRNA-n27i-5p	T_00016031		2.5	
CsmiRNA-n27j			1.5	
CsmiRNA-n27k			1.5	
CsmiRNA-n27m			1.5	
CsmiRNA-n27n-5p			1.5	
CsmiRNA-n27o			1	
CsmiRNA-n27p			2	
CsmiRNA-n56b.1	T_00016799		2	
miR5569b	T_00035890		3	
CsmiRNA-n1a-5P			3	
miR5569b	T_00035891		3	
CsmiRNA-n1a-5P			3	

Table 1. *Cont.*

miRNA Name	Target ID	Target Annotation	Expectation	Pathway
miR5569b CsmiRNA-n1a-5P	T_00035895		3 3	
CsmiRNA-n22c-5p	T_00043674		3	
CsmiRNA-n22c-5p	T_00043675		3	
CsmiRNA-n22c-5p	T_00043676		3	
CsmiRNA-n22c-5p	T_00043677		3	
CsmiRNA-n25m.1	T_00051204		3	
CsmiRNA-n37a-3p CsmiRNA-n37b-3p	T_00068566		3 3	
CsmiRNA-n37a-3p CsmiRNA-n37b-3p	T_00068567		3 3	
miR5569b CsmiRNA-n1a-5P	T_00078801		3 3	
miR5569b CsmiRNA-n1a-5P	T_00078802		3 3	
miR5569b CsmiRNA-n1a-5P	T_00078803	Kinesin-like protein	3 3	
miR5569b CsmiRNA-n1a-5P	T_00078805		3 3	
	T_00080222		3	
	T_00080223		3	
	T_00080224		3	
	T_00087551		3	
CsmiRNA-n23d-5p	T_00087552		3	
	T_00087553		3	
	T_00087554		3	
	T_00087555		3	
	T_00087556		3	
	T_00087557		3	
	T_00087558		3	
	T_00087559		3	
	T_00087560		3	
miR482a-3p miR482b-3p miR482c-3p	T_00093433		3 3 3	
miR482a-3p miR482b-3p miR482c-3p	T_00093434	CESA	3 3 3	
miR482a-3p miR482b-3p miR482c-3p	T_00093435		3 3 3	
miR1508-5P CsmiRNA-n34c-3p	T_00012561 T_00012569		3 3	
CsmiRNA-n15b-5p	T_00054754	VILLIN	3	

Table 1. *Cont.*

miRNA Name	Target ID	Target Annotation	Expectation	Pathway
miR172a-3p CsmiRNA-n4b-5p	T_00062735	FRA	3 3	
miR172a-3p CsmiRNA-n4b-5p	T_00062736		3 3	
CsmiRNA-n24m.1	T_00064946	CSL	2	
CsmiRNA-n32e.2	T_00019352	UDP-galactose transporter	2.5	
miR171a-5p miR171b-5p miR171c-5p	T_00021321		3 3 3	

AAE: acyl-activating enzyme; DXR: 1-deoxy-D-xylulose 5-phosphate reductoisomerase; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; LOX: lipoxygenase; ACCase: acetyl-CoA carboxylase; COBL: COBRA-like protein; CESA: Cellulose synthase; FRA: inositol polyphosphate 5-phosphatase; CSL: Cellulose synthase-like protein.

Subsequently, the crosstalk between AS and miRNAs involved in cannabinoids, FA and cellulose was analyzed. As shown in Table S12, there were two types for the miRNAs acting on the isoforms of the structural gene. One was miRNAs targeting all the isoforms of one structural gene. For example, CsmiRNA-n33j.1-3p targeted the transcripts (T_00033798/00033799/00033800) of AAE. The other was the miRNAs targeting part of isoforms of one structural gene. For example, there were six transcripts (T_00069671/00069672/00069673/00069674/00069675/00069676) of *AAE*, in which only the transcripts (T_00069673/00069674/00069676) were targeted by CsmiRNA-n33j.1-3p.

2.7. Cleavage Targets Guided by miRNAs

miRNAs could act on their targets by cleavage [80]. In order to validate the authentic cleavage targets of miRNAs, a degradome library was sequenced, and psRobot software was used to analyze the cleavage site of target genes and the reads of degradation fragments [81]. As a result, a total of 10,810,702 clean reads represented by 5,623,324 unique sequences were obtained after high-quality filtering and removing the adapters and null inserts. Most of the miRNAs in plants target genes at the 9th, 10th and 11th nucleotides of the complementary region between miRNA and mRNA, occasionally at the other site [82]. The cleavage sites of miRNA targets were predicted by psRobot software, as shown in Table S13.

In total, 145 transcripts were identified, which were targeted by 62 miRNAs, including 40 conserved miRNAs and 22 *C. sativa*-specific miRNAs (Table S14). The major targets of conserved miRNAs were TFs involved in development, responding to abiotic stresses and metabolism. In aspects of development, for example, miR156 targeted 34 squamosa promoter binding protein like transcripts (*SPLs*). In *Arabidopsis thaliana* L. *SPLs* played an important part in flowering development [83]. miR160 targeted six auxin response factor 16 (*ARF16*) transcripts. It was reported that miR160 regulated auxin mediated development by ARF10/16/17 in tomato through posttranscriptional regulation [84]. MiR164 targeted 10 CUP SHAPED COTYLEDON 2 (*CUC2*) transcripts. MiR164 regulated the axillary meristem formation by targeting *CUC1* and *CUC2*, participating in the development of cotyledons and floral organs in *Arabidopsis* [85]. In response to abiotic stresses, 20 transcripts encoding NAC TFs were identified as targets of miR164. NAC TFs were validated for response to salinity as well as high temperature and drought stresses in *Brassica juncea* L. Czern [86]. In aspects of metabolism, miR393 targeted two transport inhibitor response 1 (*TIR1*) genes and two auxin signaling f-box 2 (*AFB2*) genes. In *Arabidopsis*, *TIR1* and *AFB2* had an effect on the auxin sensitivity [87]. In addition, 16 ubiquitin-conjugating E2 enzyme transcripts were targeted by miR399. A previous study showed that miR399 targeted ubiquitin-conjugating E2 enzyme genes functioning in the control of inorganic phosphate homeostasis in *Arabidopsis* [88].

A few *C. sativa*-specific miRNAs were also found, including CsmiRNA-n56/n10/n27/n30 and other proteins. For catalysis, CsmiRNA-n56 targeted two cytokinin hydroxylase (*CYP735A1*) genes. CYP735A1 is a cytochrome P450 monooxygenase, and it was demon-

strated that it catalyzed the biosynthesis of *trans*-Zeatin in *Arabidopsis* [89]. CsmiRNA-n10 targeted the 3-hydroxy-3-methylglutaryl-coenzyme A reductase gene, which catalyzed the HMG-CoA to form MVA [90]. CsmiRNA-n27 targeted the U-box domain-containing protein, which participated in the initial fruit developmental stage in *Musa acuminate* L [91]. In addition, CsmiRNA-n30 targeted tubby-like F-box protein, which functioned in abiotic stress tolerance in *Manihot esculenta* Crantz [92]. Furthermore, CsmiRNA-n30 also targeted *SURP* and G-patch domain-containing protein 1-like protein (*SUGP1*). SURP and SUGP1 were associated with plasma LDL cholesterol levels, coronary artery disease and some other energy metabolism [93].

2.8. Identification of lncRNAs in C. sativa

To decipher the functions of lncRNAs in *C. sativa*, a pipeline was firstly established to predict the lncRNAs. After removing 48,800 transcripts without Nr annotations, the 1375 transcripts less than 300 bp in length were discarded. Then, the remaining transcripts were filtered by ORF $\leq$ 100 aa using ESTScan, leaving 8260 genes. Next, the CPAT was used to filter the 430 possible protein-coding genes, leaving a total of 7830 transcripts. Finally, BLASTN was used to remove the housekeeping npcRNAs (47) from the alignment with the Rfam database, thereby obtaining 7783 lncRNA candidates and 47 housekeeping npcRNAs (Table S15, Figure 4A).
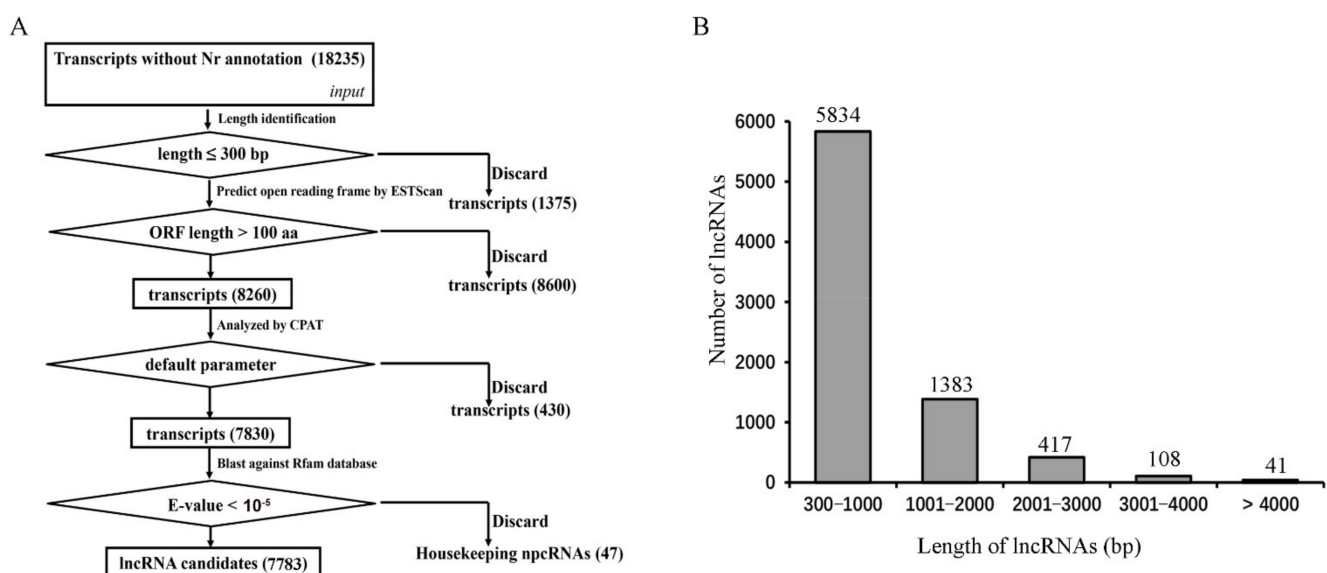

Figure 4. The schematic pipeline and size distribution of lncRNAs. (**A**) Predicted pipeline of lncRNA candidates. (**B**) Size distribution of lncRNA candidates.

After that, an assay of the length distribution of lncRNAs was conducted. It was found that 5834 lncRNAs ranged from 300 to 1000 bp. Most of the lncRNAs were distributed between 1000 and 2000 bp. The others included 525 lncRNA candidates with 2000–4000 bp and 41 over 4000 bp in length (Figure 4B).

2.9. Target Prediction for lncRNAs in C. sativa

BLASTN was used to predict the targets of lncRNAs to reveal their regulatory role in *C. sativa*. Among the 7783 lncRNAs, 1516 lncRNAs were revealed to target 3452 transcripts (Table S16). Among the 3452 transcripts, 1536 were annotated by the KEGG database. KEGG enrichment showed that 26, 25, 102, 224 and 25 genes were involved in cellular processes, environmental information processing, genetic information processing, metabolism and organismal systems, respectively, in which the top five enrichment pathways were 1,4-alpha-glucan branching enzyme, uroporphyrinogen-III synthase, RecQ-

mediated genome instability protein 1, structure-specific endonuclease subunit SLX1 and flap endonuclease-1 (Figure S2).

Then, the targets related to the biosynthetic pathway of cannabinoids, FA and cellulose were focused on. Therefore, lncTar software was used to further predict the interactions between lncRNAs and target genes. As a result, nine lncRNAs were found to target the *LOX, GPPS, AAE* and *DXS* genes in the cannabinoid biosynthetic pathway, and eight lncRNAs were found to target the *COBL*, kinesin-like and *MAPKKK* genes in the cellulose metabolism (Table 2). In addition, 111 TFs, 59 TRs and 141 PKs were also identified, which were targeted by 92, 90 and 135 lncRNAs, respectively (Table S17).

Subsequently, the crosstalk between AS and lncRNAs related to the cannabinoids and cellulose were analyzed. For lncRNAs, on the one hand, all of the isoforms (T_00013932 and T_00013933) targeted all the kinesin-like protein gene isoforms (T_00013935/13936/13937). On the other hand, part isoforms of lncRNAs (T_00056848/56849/56852/56853/56854 and T_00020067) targeted one transcript of GPPS (T_00049365) and DXS (T_00064406). For target genes, all of the *LOX* isoforms (T_00022090 and T_00022091) were targeted by different lncRNAs (T_00010712 and T_00019445), and parts of AS of kinesin-like protein gene isoforms (T_00043674/43675/43677/43678/43681) were targeted by different lncRNAs (T_00043682/29565/77089/43682), respectively (Table S18).

Table 2. LncRNAs targeted structural genes potentially involved in cannabinoids and cellulose biosynthesis.

LncRNA ID	Target ID	Target Annotation	Start Position LncRNA	End Position LncRNA	Start Position Target	End Position Target	Pathway
T_00010712	T_00010715		1	228	1261	1488	
T_00010712	T_00022090	LOX	1	228	1947	2174	
T_00019445			127	556	1	430	
T_00019445	T_00022091		127	556	1	430	
T_00056848	T_00049365	GPPS	1	164	1784	1947	cannabinoids
T_00056849			1	164	1784	1947	
T_00056852			1	162	1786	1947	
T_00056853			1	161	1787	1947	
T_00056854			1	153	1795	1947	
T_00090880	T_00058852	AAE	1	369	19	387	
T_00020067	T_00064406	DXS	383	863	1	481	
T_00013932	T_00013935		1	1075	1	1075	
T_00013933			1	1066	10	1075	
T_00013932	T_00013936		1	1079	1	1079	
T_00013933			1	1070	10	1079	
T_00013932	T_00013937		1	1077	1	1077	
T_00013933		Kinesin-like protein	1	1068	10	1077	
T_00052635	T_00016799		558	982	1	425	cellulose
T_00043682	T_00043674		174	1620	1	1447	
T_00029565	T_00043675		357	542	1	186	
T_00077089			501	638	1	138	
T_00043682	T_00043677		174	1620	1	1447	
T_00043682	T_00043678		174	1620	1	1447	
T_00043682	T_00043681		174	1620	1	1447	
T_00030548	T_00016031	COBL	1	346	957	1302	
T_00059592	T_00061939	MAPKKK	1	1031	2317	3347	

GPPS: Geranyl diphosphate synthase; MAPKKK: mitogen-activated protein kinase kinase.

2.10. QRT-PCR Analysis of the Expression Files of lncRNAs and Target Genes

Seven lncRNA/structural gene pairs were selected for qRT-PCR validation, and the primer sequences are shown in Table S19. As a result, three lncRNA/structural gene pairs were identified with negative expression files in at least four tissues, including lncR880 (T_00090880)/*AAE* (T_00058852), lncR682 (T_00043682)/kinesin-like protein (T_00043674) and lncR578 (T_00030548)/*COBL* (T_00016031) (Figure 5), indicating the negative regulation of these lncRNAs.
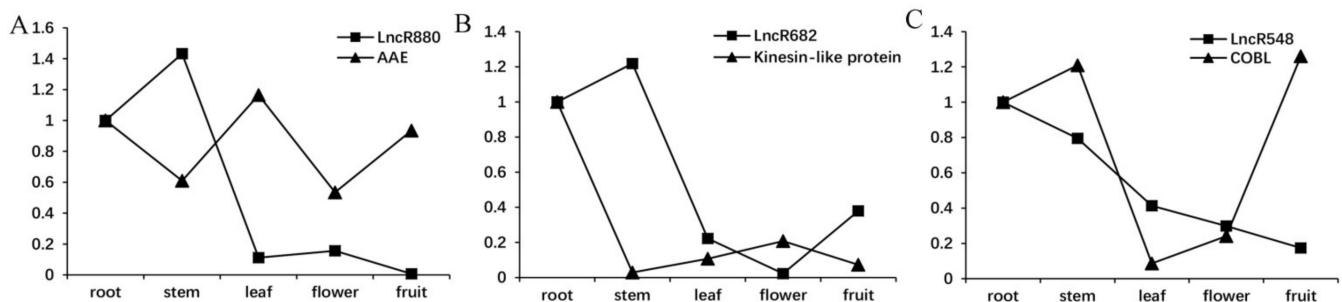
Figure 5. QRT-PCR validation of expression profiles of lncRNA/structural gene pairs of lncR880/AAE (**A**), lncR682/kinesin-like protein (**B**) and lncR578/COBL (**C**). The average expression profile of every gene in the root was chosen as a control and set to 1.

3. Discussion

In this study, transcriptome, small RNA and degradome libraries of hemp 'Yunma No.1' were constructed and comprehensively analyzed. Compared with previous studies in *C. sativa*, three advancements were made. Firstly, the protein-coding genes related to cannabinoids, FA and cellulose were mined, and the transporter and regulatory protein genes were systematically predicted.

Secondly, to reveal the gene expression and regulation at post-transcriptional level, the AS, miRNAs and lncRNAs in *C. sativa* firstly were predicted. A total of 27,907 AS events in *C. sativa* were identified, and five AS events were validated by RT-PCR at random. Among them, AS occurred in some structural genes of the cannabinoids, FA and cellulose biosynthetic pathway.

MiRNAs and lncRNAs functioned in multiple pathways. We focused on the top five KEGG enrichment pathways of miRNA/lncRNA targets. For miRNAs, it was reported that miR-362 targeted RecQ-mediated genome instability protein 1 [94]. miR-125b targeted ER alpha-1, 2-mannosidase and played a critical role in maintaining protein homeostasis in the mammalian secretory pathway [95], while, KAS II, GPI ethanolamine phosphate transferase 3 subunit O and CCR4-NOT transcription complex subunit 4 potentially targeted by miRNAs were not reported. KAS II was involved in fatty acid elongation from 16:0-ACP to 18:0-ACP [96]. GPI ethanolamine phosphate transferase 3 subunit O functioned in GPI-anchor biosynthesis [97]. The top five KEGG enrichment pathways targeting lncRNAs have not been revealed to date. For example, 1,4-alpha-glucan branching enzyme played a significant role in the biosynthesis of branched polysaccharides, glycogen and amylopectin [98]. All these results indicate the important regulation of miRNAs and lncRNAs in *C. sativa*.

MiRNAs and lncRNAs potentially regulate the biosynthesis of cannabinoids, FA and cellulose in two aspects. On the one hand, they directly target the structural genes (Tables 1 and 2, Figure 5). On the other hand, they targeted the structural gene related regulatory protein genes. Previously, we determined some GA- or isoflavone-related TFs, TRs and PKs potentially targeted by miRNAs, indicting the significant roles of regulatory protein genes in metabolism [77,99]. Based on the known biosynthetic pathways of cannabinoids, Luo et al. successfully biosynthesized cannabinoids along with the unnatural analogues in yeast [100], but the different isoforms of structural genes, regulatory protein genes and non-coding RNAs were not considered. In this study, only a strand-specific

RNA-Seq library was constructed. In the future, more libraries should be sequenced, and co-expression analysis will be carried out between the interested structural genes and regulatory genes to screen out and validate highly relevant regulators and make full use of them.

Additionally, the crosstalk between AS and miRNAs or lncRNAs in the biosynthesis of cannabinoids, FA and cellulose was studied for the first time in *C. sativa* (Tables S12 and S18). AS could regulate miRNA and lncRNA-mediated gene regulation by producing mRNA isoforms, which contained miRNAs/non-coding RNAs target sites or not [101]. A previous study revealed that the 44 of 354 identified miRNA binding sites were affected by AS in *Arabidopsis* [42]. Some lncRNA genes also possessed introns, leading to the production of lncRNA isoforms [102]. In *Arabidopsis*, Calixto et al. identified 135 cold-responsive lncRNAs, of which AS occurred in one-third [103]. In this research, the crosstalk between 17 lncRNAs and their target transcripts was predicted in *C. sativa* (Table S18).

During our work, a PacBio SMRT and a nanopore-based assembly genome of *C. sativa* was released in the NCBI under the accession numbers PRJNA73819 and GCA_900626175.2, providing good sources for transcriptome assembly of *C. sativa* for the future [16,17].

Taken together, all these results enhance the knowledge of gene expression and regulation in *C. sativa*.

4. Materials and Methods

4.1. Plant Materials and RNA Extraction

Hemp 'Yunma No.1' plants were cultivated in field conditions at the Yuanjiang experimental station of Bast Fiber Crops, Chinese Academy of Agricultural Sciences. Different tissues including flowers, fruits, stems, leaves and roots were collected from three female hemp plants in September 2014 and quickly frozen in liquid nitrogen. Total RNAs of fruits and roots were extracted using Trizol Reagent (Invitrogen, Carisbad, CA, USA). Total RNAs of flowers, stems and leave were extracted using Lysis Buffer PL (Bioteke, Wuxi, China). Then, these RNAs were treated by RQ1 DNase (Promega, Madison, WI, USA) to eliminate the genome DNA. Briefly, about 100 μg total RNA of each sample was incubated with 20 U DNase I at 37 °C for 30 min. Then, equal volumes of phenol/chloroform/isoamyl alcohol (25:24:1) were added and mixed. After centrifugation at 12,000 rpm for 5 min at room temperature, the upper layer was transferred to a new tube. Subsequently, 1/10 volume of 3 M sodium acetate and 2 volumes of chilled ethanol were added, mixed and kept at for 20 min at −80 °C. Finally, after centrifugation at 12,000 rpm for 10 min at 4 °C, the precipitation was washed with chilled 70% ethanol, dried and dissolved in a suitable amount of DEPC-treated water. Equal quantities of total RNAs from different tissues were pooled together, then their integrity and concentrations were examined using a NanoDrop 2000c bioanalyzer (Thermo Fisher Scientific, Waltham, MA, USA). The pooled RNA with an integrity number of more than seven was used for the following RNA-Seq, small RNA and degradome library construction.

4.2. RNA-Seq Library Construction and Analysis

The RNA-Seq library was built by using the dUTP method according to the manufacturer's protocol [104]. First, mRNA was isolated from 2 μg total RNA by using Oligo (dT) magnetic beads. Then, about 20 ng of mRNA was used for library construction. After the RNA sample was qualified, the mRNA was enriched with magnetic beads with Oligo (dT). Fragmentation buffer was used for breaking mRNA into short fragments, which were used as a template to synthesize one-strand cDNA with six base random primers (random hexamers). Then, we added buffer, dNTPs (the dTTP in dNTP was replaced by dUTP) and DNA Polymerase I and RNase H to synthesize two-strand cDNA. The double-strand cDNA was purified by AMPure XP beads. The second strand of cDNAs containing Us was degraded by USER enzyme and further repaired, A-tailed and connected to the sequencing adapter. Finally, PCR amplification was performed, and the final library was purified by AMPure XP beads. After the library was constructed, Qubit was used for

preliminary quantification, and then Agilent 2100 was used to detect the size of the insert in the library. High-throughput sequencing was carried out by using a HiSeq 2000 platform. Raw reads were first processed using Trimmomatic with the following options: ILLUMINACLIP: TruSeq3-PE. fa: 2:30:10 LEADING:5 TRAILING:5 to remove adapters and discard reads containing low quality or N bases (below quality 5). After filtering, the clean reads were aligned with the *C. sativa* 'Purple Kush' genome [18] using tophat with the -a/−min-anchor 8 option, which required that the anchor length should be more than 8bp for splicing-junction reads [18,44]. Then, the mapped reads were assembled by Cufflinks software using -F 0.05 -A 0.01 -I 15000 [45]. Subsequently, the assembled files were used to identify the AS events and types by ASTALAVISTA with default options using the following command: /usr/java/jdk1.6.0_45/jre/bin/java -Xmx4G -jar/srv/web/cgi -bin/astal avista/jar/gphase5.jar astalavista.gtf -nonmd -genome csa -output gtf -html -html 2 > &1 [73]. To determine gene functions, the assembly transcripts were annotated by NR and KEGG databases.

4.3. Validation of AS through Reverse Transcriptional PCR

Total RNA (2 μg) as described for RNA-Seq library construction was used for reverse transcription PCR (RT-PCR) to validate the AS events. Briefly, RT was carried out using 2 μg of total RNA by 200 U M-MLV Reverse Transcriptase (TaKaRa, Dalian, China) in a 20 μL volume under the following the conditions: 65 °C for 5 min; 25 °C for 10 min; 42 °C for 60 min and 70 °C for 15 min. The resulting cDNA was used for PCR amplification, which was carried out under the following conditions: 95 °C for 3 min; 45 cycles of 94 °C for 30 s; 58 °C for 30 s and 72 °C for 15 s. The PCR products were separated by electrophoresis with a 3% agarose gel. The primers are listed in Table S19.

4.4. Small RNA Library Construction and Analysis

The small RNA library was constructed as described [105]. Briefly, 10–30-nt small RNAs were purified from 1 μg Total RNA by a 15% denaturing polyacrylamide gel and then ligated to the adapters. After reverse transcription by M-MLV (TaKaRa, Dalian, China), these small RNAs were amplified by PCR and isolated by 6% polyacrylamide gel, and suitable sizes were recovered and checked by qPCR. Finally, the library was sequenced by the HiSeq 2000 platform at Beijing HiTseq Technology Co. Ltd. (Beijing, China).

After sequencing, the raw reads were first processed using Trimmomatic software [106] with the option LEADING:5 TRAILING:5 to discard reads containing low quality or N bases (below quality 5). Then, adaptor and shorter reads (less than 16 nt) were removed by fastx_clipper in the FASTX-Toolkit v0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/) (2 February 2010) using the option fastx_clipper -Q 33 -a CTGTAGGCACCATCAATCA -l 16 -d 0 -n -v -M 4. Then, the clean small RNAs with counts ≥5 ranging from 18 to 30 nt were further analyzed. These small RNAs were compared with the *C. sativa* 'Purple Kush' genome sequences by the psRobot software (v1.2) [81] and mireap software [107] using the default parameters. Then, these predicted miRNAs were checked manually according to the criterion of miRNAs [108]. Minimum free energy index (MFEI) was also calculated [109], and a cutoff >0.85 was used to distinguish miRNAs from protein-coding genes because the MFEI of more than 90% of miRNA precursors is >0.85.

4.5. Prediction of MiRNA Targets

Through the psRNATarget Server using an expectation penalty score of ≤3, targets of miRNAs were predicted [110].

4.6. Degradome Library Construction, Sequencing and Analysis

Briefly, about 150 μg pooled total RNA was used for degradome library construction by BGI (Shenzhen, China) following a published protocol [111]. In brief, mRNAs were isolated using oligo(dT) magnetic beads. Then, single-stranded 5′ RNA adaptors were ligated to mRNA fragments using T4 RNA ligase (Ambion, Austin, TX, USA) and

then reverse-transcribed into cDNA using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA). After digestion with MmeI, 3′ DNA adaptors were ligated to the digested DNA fragments. Finally, the products were amplified by PCR and sequenced using the HiSeq 2000 system (BGI, Shenzhen, China). The cleavage sites and the targets of miRNAs were identified using the psRobot software (v1.2) using the options -ts 2.5 -fp 2 -tp 17 -gl 17 -p 1 -gn 1 [81].

4.7. Prediction of LncRNAs

A pipeline was established for identification of lncRNAs in hemp (Figure 4). First, the assembly transcripts without Nr annotation were further analyzed. Then, the transcripts less than 300 base pairs (bp) in length were removed. Subsequently, the remaining transcripts were filtered by ESTScan with a cutoff of ORF $\geq$ 100 amino acids (aa) [112]. Next, the coding potential assessment tool (CPAT) was used to filter the possible protein coding genes [113]. Finally, BLASTN program was used to remove the housekeeping npcRNAs annotated by the Rfam database.

4.8. Prediction of LncRNA Targets

Targets of lncRNAs were predicted by using BLASTN to search the lncRNAs with regions significantly complementary to the protein-coding genes using the cutoff 1×10^{-5} [114,115]. Then, the interesting lncRNA-RNA interactions were further analyzed by LncTar using the cutoff of the normalized deltaG (ndG) $\leq$ 0.1 [116].

4.9. QRT-PCR Analysis of the Expression Files of LncRNAs and Targets

RNA purity, concentration and integrity of each sample treated by DNase I were examined using a NanoDrop 2000c bioanalyzer (Thermo Fisher Scientific, Waltham, MA, USA). Then, RT was carried out using 1 µg total RNA and 200 U M-MLV Transcriptase (TaKaRa, Dalian, China) in a 20 µL volume. The reaction was conducted at 70 °C for 10 min, 42 °C for 60 min and 70 °C for 15 min. The RT product was diluted 40 times with sterile water. The qPCR was conducted on the BIO-RAD CFX system in a 20 µL volume containing 4 µL diluted cDNA, 250 nM forward primer, 250 nM reverse primer, and 1 × SYBR Premix Ex Taq II (TaKaRa, Dalian, China) using the following conditions: 95 °C for 3 min, 40 cycles of 95 °C for 15 s, 59 °C for 15 s and 72 °C for 15 s. Each test was performed in triplicate. The 40S ribosomal protein S8 (T_00004056) was used as an internal reference because its expression level is relatively stable in different tissues. Melting curves were analyzed to verify the specificity using the Bio-Rad CFX Manager software. The relative expression levels were calculated using the $2^{-\Delta\Delta CT}$ method [117]. The primer sequences are listed in Table S20.

References

1. Chen, T.; He, J.; Zhang, J.; Zhang, H.; Qian, P.; Hao, J.; Li, L. Analytical characterization of Hempseed (seed of *Cannabis sativa* L.) oil from eight regions in China. *J. Diet. Suppl.* **2010**, *7*, 117–129. [CrossRef] [PubMed]
2. Aiello, G.; Fasoli, E.; Boschin, G.; Lammi, C.; Zanoni, C.; Citterio, A.; Arnoldi, A. Proteomic characterization of hempseed (*Cannabis sativa* L.). *J. Proteom.* **2016**, *147*, 187–196. [CrossRef] [PubMed]
3. Paredez, A.R.; Somerville, C.R.; Ehrhardt, D.W. Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* **2006**, *312*, 1491–1495. [CrossRef] [PubMed]
4. Potter, D.J.; Clark, P.; Brown, M.B. Potency of delta 9-THC and other cannabinoids in cannabis in England in 2005: Implications for psychoactivity and pharmacology. *J. Forensic Sci.* **2008**, *53*, 90–94. [CrossRef] [PubMed]
5. Mehmedic, Z.; Chandra, S.; Slade, D.; Denham, H.; Foster, S.; Patel, A.S.; Ross, S.A.; Khan, I.A.; ElSohly, M.A. Potency trends of Δ9-THC and other cannabinoids in confiscated cannabis preparations from 1993 to 2008. *J. Forensic Sci.* **2010**, *55*, 1209–1217. [CrossRef] [PubMed]
6. Booth, J.K.; Page, J.E.; Bohlmann, J. Terpene synthases from *Cannabis sativa*. *PLoS ONE* **2017**, *12*, e0173911. [CrossRef] [PubMed]
7. Bonini, S.A.; Premoli, M.; Tambaro, S.; Kumar, A.; Maccarinelli, G.; Memo, M.; Mastinu, A. *Cannabis sativa*: A comprehensive ethnopharmacological review of a medicinal plant with a long history. *J. Ethnopharmacol.* **2018**, *227*, 300–315. [CrossRef]
8. Grof, C.P.L. Cannabis, from plant to pill. *Br. J. Clin. Pharmacol.* **2018**, *84*, 2463–2467. [CrossRef]
9. Lastres-Becker, I.; Molina-Holgado, F.; Ramos, J.A.; Mechoulam, R.; Fernandez-Ruiz, J. Cannabinoids provide neuroprotection against 6-hydroxydopamine toxicity in vivo and in vitro: Relevance to Parkinson's disease. *Neurobiol. Dis.* **2005**, *19*, 96–107. [CrossRef]
10. Mechoulam, R.; Peters, M.; Murillo-Rodriguez, E.; Hanus, L.O. Cannabidiol—Recent advances. *Chem. Biodivers.* **2007**, *4*, 1678–1692. [CrossRef]
11. Gaoni, Y.; Mechoulam, R. Isolation, Structure, and Partial Synthesis of an Active Constituent of Hashish. *J. Am. Chem. Soc.* **1964**, *86*, 1646–1647. [CrossRef]
12. Yotoriyama, M.; Numao, K. Phenotype of cannabinoids obtained by crossing cannabidiolic acid (CBDA)-cannabis and tetrahydro-cannabinolic acid (THCA)-cannabis. *Jpn. J. Toxicol. Environ. Health* **1979**, *25*, 166–168. [CrossRef]
13. Goldberg, E.; Gakhar, N.; Ryland, D.; Aliani, M.; Gibson, R.; House, J. Fatty Acid Profile and Sensory Characteristics of Table Eggs from Laying Hens Fed Hempseed and Hempseed Oil. *J. Food Sci.* **2012**, *77*, S153–S160. [CrossRef]
14. Simopoulos, A.P. The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomed. Pharmacother.* **2002**, *56*, 365–379. [CrossRef]
15. Shen, C.; Li, X.; Tian, K.; Zhang, B.; Chen, Q. Experimental analysis on mechanical model of ramie stalk. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 26–33.
16. Grassa, C.J.; Weiblen, G.D.; Wenger, J.P.; Dabney, C.; Poplawski, S.G.; Timothy Motley, S.; Michael, T.P.; Schwartz, C.J. A new Cannabis genome assembly associates elevated cannabidiol (CBD) with hemp introgressed into marijuana. *New Phytol.* **2021**, *230*, 1665–1679. [CrossRef]
17. Laverty, K.U.; Stout, J.M.; Sullivan, M.J.; Shah, H.; Gill, N.; Holbrook, L.; Deikus, G.; Sebra, R.; Hughes, T.R.; Page, J.E.; et al. A physical and genetic map of *Cannabis sativa* identifies extensive rearrangements at the *THC/CBD acid synthase* loci. *Genome Res.* **2019**, *29*, 146–156. [CrossRef]
18. Van Bakel, H.; Stout, J.M.; Cote, A.G.; Tallon, C.M.; Sharpe, A.G.; Hughes, T.R.; Page, J.E. The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* **2011**, *12*, R102. [CrossRef]
19. Yin, D.-D.; Li, S.-S.; Shu, Q.-Y.; Gu, Z.-Y.; Wu, Q.; Feng, C.-Y.; Xu, W.-Z.; Wang, L.-S. Identification of microRNAs and long non-coding RNAs involved in fatty acid biosynthesis in tree peony seeds. *Gene* **2018**, *666*, 72–82. [CrossRef]
20. Liao, B.; Hao, Y.; Lu, J.; Bai, H.; Guan, L.; Zhang, T. Transcriptomic analysis of *Perilla frutescens* seed to insight into the biosynthesis and metabolic of unsaturated fatty acids. *BMC Genom.* **2018**, *19*, 213. [CrossRef]
21. Meents, M.J.; Watanabe, Y.; Samuels, A.L. The cell biology of secondary cell wall biosynthesis. *Ann. Bot.* **2018**, *121*, 1107–1125. [CrossRef]
22. Kim, H.K.; Pham, M.H.C.; Ko, K.S.; Rhee, B.D.; Han, J. Alternative splicing isoforms in health and disease. *Pflugers Arch. Eur. J. Physiol.* **2018**, *470*, 995–1016. [CrossRef]
23. Sun, Y.; Zhang, Q.; Liu, B.; Lin, K.; Zhang, Z.; Pang, E. CuAS: A database of annotated transcripts generated by alternative splicing in cucumbers. *BMC Plant Biol.* **2020**, *20*, 119. [CrossRef]
24. Sun, H.X.; Chua, N.H. Bioinformatics Approaches to Studying Plant Long Noncoding RNAs (lncRNAs): Identification and Functional Interpretation of lncRNAs from RNA-Seq Data Sets. *Methods Mol. Biol.* **2019**, *1933*, 197–205. [CrossRef]
25. Tafrihi, M.; Hasheminasab, E. MiRNAs: Biology, Biogenesis, their Web-based Tools, and Databases. *MicroRNA* **2019**, *8*, 4–27. [CrossRef]
26. Bartel, D.P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* **2004**, *116*, 281–297. [CrossRef]
27. Brodersen, P.; Sakvarelidze-Achard, L.; Bruun-Rasmussen, M.; Dunoyer, P.; Yamamoto, Y.Y.; Sieburth, L.; Voinnet, O. Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **2008**, *320*, 1185–1190. [CrossRef]
28. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001**, *411*, 494–498. [CrossRef]

29. Hamilton, A.J.; Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **1999**, *286*, 950–952. [CrossRef]

30. Carrington, J.C.; Ambros, V. Role of microRNAs in plant and animal development. *Science* **2003**, *301*, 336–338. [CrossRef]

31. Hake, S. MicroRNAs: A role in plant development. *Curr. Biol.* **2003**, *13*, R851–R852. [CrossRef] [PubMed]

32. Kidner, C.A.; Martienssen, R.A. The developmental role of microRNA in plants. *Curr. Opin. Plant Biol.* **2005**, *8*, 38–44. [CrossRef] [PubMed]

33. Ma, Y.; Cao, J.; He, J.; Chen, Q.; Li, X.; Yang, Y. Molecular Mechanism for the Regulation of ABA Homeostasis During Plant Development and Stress Responses. *Int. J. Mol. Sci.* **2018**, *19*, 3643. [CrossRef] [PubMed]

34. Ku, Y.S.; Wong, J.W.; Mui, Z.; Liu, X.; Hui, J.H.; Chan, T.F.; Lam, H.M. Small RNAs in Plant Responses to Abiotic Stresses: Regulatory Roles and Study Methods. *Int. J. Mol. Sci.* **2015**, *16*, 24532–24554. [CrossRef]

35. Gou, J.Y.; Felippes, F.F.; Liu, C.J.; Weigel, D.; Wang, J.W. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* **2011**, *23*, 1512–1522. [CrossRef]

36. Lu, S.; Li, Q.; Wei, H.; Chang, M.J.; Tunlaya-Anukit, S.; Kim, H.; Liu, J.; Song, J.; Sun, Y.H.; Yuan, L.; et al. Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10848–10853. [CrossRef]

37. Ng, D.W.; Zhang, C.; Miller, M.; Palmer, G.; Whiteley, M.; Tholl, D.; Chen, Z.J. cis- and trans-Regulation of miR163 and target genes confers natural variation of secondary metabolites in two Arabidopsis species and their allopolyploids. *Plant Cell* **2011**, *23*, 1729–1740. [CrossRef]

38. Beltran, M.; Puig, I.; Peña, C.; García, J.M.; Alvarez, A.B.; Peña, R.; Bonilla, F.; de Herreros, A.G. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev.* **2008**, *22*, 756–769. [CrossRef]

39. Guo, L.; Zhao, Y.; Yang, S.; Zhang, H.; Chen, F. An integrated analysis of miRNA, lncRNA, and mRNA expression profiles. *BioMed Res. Int.* **2014**, *2014*, 345605. [CrossRef]

40. Melamed, Z.; Levy, A.; Ashwal-Fluss, R.; Lev-Maor, G.; Mekahel, K.; Atias, N.; Gilad, S.; Sharan, R.; Levy, C.; Kadener, S.; et al. Alternative splicing regulates biogenesis of miRNAs located across exon-intron junctions. *Mol. Cell* **2013**, *50*, 869–881. [CrossRef]

41. Jia, F.; Rock, C.D. *MIR846* and *MIR842* comprise a cistronic *MIRNA* pair that is regulated by abscisic acid by alternative splicing in roots of Arabidopsis. *Plant Mol. Biol.* **2013**, *81*, 447–460. [CrossRef]

42. Yang, X.; Zhang, H.; Li, L. Alternative mRNA processing increases the complexity of microRNA-based gene regulation in *Arabidopsis*. *Plant J.* **2012**, *70*, 421–431. [CrossRef]

43. Xue-Li, H.U.; Ming, Y.; Yan-Ping, X.U.; Meng-Bi, G.; Yu, C.; Wen-Pin, D.; Ju-Xian, W.U.; Min, C.; Hong-Yan, G. Effect of Planting Density on Yield and Agronomic Characters of Hemp Variety Yunma No.1. *Plant Fiber Sci. China* **2009**, *31*, 322–324+303.

44. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **2013**, *14*, R36. [CrossRef]

45. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [CrossRef]

46. Yangyang, D.; Jianqi, L.I.; Songfeng, W.U.; Yunping, Z.H.U.; Yaowen, C.; Fuchu, H.E. Integrated nr database in protein annotation system and its localization. *Comput. Eng.* **2006**, *32*, 71–72.

47. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [CrossRef]

48. Cahoon, E.B.; Kinney, A.J. Dimorphecolic acid is synthesized by the coordinate activities of two divergent Delta12-oleic acid desaturases. *J. Biol. Chem.* **2004**, *279*, 12495–12502. [CrossRef]

49. Rea, K.A.; Casaretto, J.A.; Al-Abdul-Wahid, M.S.; Sukumaran, A.; Geddes-McAlister, J.; Rothstein, S.J.; Akhtar, T.A. Biosynthesis of cannflavins A and B from *Cannabis sativa* L. *Phytochemistry* **2019**, *164*, 162–171. [CrossRef]

50. Torres-Franklin, M.-L.; Repellin, A.; Biet, H.; d'Arcy-Lameta, A.; Zuily-Fodil, Y.; Pham-Thi, A.-T. Omega-3 fatty acid desaturase (FAD3, FAD7, FAD8) gene expression and linolenic acid content in cowpea leaves submitted to drought and after rehydration. *Environ. Exp. Bot.* **2009**, *65*, 162–169. [CrossRef]

51. Falcone, D.L.; Gibson, S.; Lemieux, B.; Somerville, C. Identification of a gene that complements an *Arabidopsis* mutant deficient in chloroplast omega 6 desaturase activity. *Plant Physiol.* **1994**, *106*, 1453–1459. [CrossRef]

52. Boesche, K.E.; Donkin, S.S. Pretreatment with saturated and unsaturated fatty acids regulates fatty acid oxidation in Madin-Darby bovine kidney cells. *J. Dairy Sci.* **2020**, *103*, 8841–8852. [CrossRef]

53. Wang, X.; Xu, R.; Wang, R.; Liu, A. Transcriptome analysis of Sacha Inchi (*Plukenetia volubilis* L.) seeds at two developmental stages. *BMC Genom.* **2012**, *13*, 716. [CrossRef]

54. Murphy, M.G. Effects of exogenous linoleic acid on fatty acid composition, receptor-mediated cAMP formation, and transport functions in rat astrocytes in primary culture. *Neurochem. Res.* **1995**, *20*, 1365–1375. [CrossRef]

55. Buchhaupt, M.; Guder, J.C.; Etschmann, M.M.; Schrader, J. Synthesis of green note aroma compounds by biotransformation of fatty acids using yeast cells coexpressing lipoxygenase and hydroperoxide lyase. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 159–168. [CrossRef]

56.　Kazaz, S.; Barthole, G.; Domergue, F.; Ettaki, H.; To, A.; Vasselon, D.; De Vos, D.; Belcram, K.; Lepiniec, L.; Baud, S. Differential Activation of Partially Redundant Δ9 Stearoyl-ACP Desaturase Genes Is Critical for Omega-9 Monounsaturated Fatty Acid Biosynthesis During Seed Development in *Arabidopsis*. *Plant Cell* **2020**, *32*, 3613–3637. [CrossRef]

57.　McFarlane, H.; Döring, A.; Persson, S. The Cell Biology of Cellulose Synthesis. *Annu. Rev. Plant Biol.* **2014**, *65*, 69–94. [CrossRef]

58.　Wei, Z.; Qu, Z.; Zhang, L.; Zhao, S.; Bi, Z.; Ji, X.; Wang, X.; Wei, H. Overexpression of Poplar Xylem Sucrose Synthase in Tobacco Leads to a Thickened Cell Wall and Increased Height. *PLoS ONE* **2015**, *10*, e0120669. [CrossRef]

59.　Seino, J.; Ishii, K.; Nakano, T.; Ishida, N.; Tsujimoto, M.; Hashimoto, Y.; Takashima, S. Characterization of rice nucleotide sugar transporters capable of transporting UDP-galactose and UDP-glucose. *J. Biochem.* **2010**, *148*, 35–46. [CrossRef]

60.　Song, X.; Zhang, B.; Zhou, Y. Golgi-localized UDP-glucose transporter is required for cell wall integrity in rice. *Plant Signal. Behav.* **2014**, *6*, 1097–1100. [CrossRef]

61.　Tamura, K.; Nakatani, K.; Mitsui, H.; Ohashi, Y.; Takahashi, H. Characterization of *katD*, a kinesin-like protein gene specifically expressed in floral tissues of *Arabidopsis thaliana*. *Gene* **1999**, *230*, 23–32. [CrossRef]

62.　Zhang, M.; Zhang, B.; Qian, Q.; Yu, Y.; Li, R.; Junwen, Z.; Liu, X.; Zeng, D.; Li, J.; Zhou, Y. Brittle Culm 12, a dual-targeting kinesin-4 protein, controls cell-cycle progression and wall properties in rice. *Plant J.* **2010**, *63*, 312–328. [CrossRef] [PubMed]

63.　Burk, D.H.; Liu, B.; Zhong, R.; Morrison, W.H.; Ye, Z.H. A katanin-like protein regulates normal cell wall biosynthesis and cell elongation. *Plant Cell* **2001**, *13*, 807–827. [CrossRef] [PubMed]

64.　Gillmor, C.S.; Poindexter, P.; Lorieau, J.; Palcic, M.M.; Somerville, C. Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. *J. Cell Biol.* **2002**, *156*, 1003–1013. [CrossRef]

65.　Zhong, R.; Burk, D.H.; Morrison, W.H., III; Ye, Z.H. *FRAGILE FIBER3*, an Arabidopsis gene encoding a type II inositol polyphosphate 5-phosphatase, is required for secondary wall synthesis and actin organization in fiber cells. *Plant Cell* **2004**, *16*, 3242–3259. [CrossRef]

66.　Maloney, V.J.; Mansfield, S.D. Characterization and varied expression of a membrane-bound endo-β-1,4-glucanase in hybrid poplar. *Plant Biotechnol. J.* **2010**, *8*, 294–307. [CrossRef]

67.　Bao, C.; Wang, J.; Zhang, R.; Zhang, B.; Zhang, H.; Zhou, Y.; Huang, S. Arabidopsis *VILLIN2* and *VILLIN3* act redundantly in sclerenchyma development via bundling of actin filaments. *Plant J.* **2012**, *71*, 962–975. [CrossRef]

68.　Maloney, V.J.; Samuels, A.L.; Mansfield, S.D. The endo-1,4-β-glucanase Korrigan exhibits functional conservation between gymnosperms and angiosperms and is required for proper cell wall formation in gymnosperms. *New Phytol.* **2012**, *193*, 1076–1087. [CrossRef]

69.　Wu, B.; Zhang, B.; Dai, Y.; Zhang, L.; Shang-Guan, K.; Peng, Y.; Zhou, Y.; Zhu, Z. *Brittle culm15* encodes a membrane-associated chitinase-like protein required for cellulose biosynthesis in rice. *Plant Physiol.* **2012**, *159*, 1440–1452. [CrossRef]

70.　Liu, L.; Shangguan, K.; Zhang, B.; Liu, X.; Yan, M.; Zhang, L.; Shi, Y.; Zhang, M.; Qian, Q.; Li, J.; et al. Brittle Culm1, a COBRA-Like Protein, Functions in Cellulose Assembly through Binding Cellulose Microfibrils. *PLoS Genet.* **2013**, *9*, e1003704. [CrossRef]

71.　Kalluri, U.C.; Payyavula, R.S.; Labbe, J.L.; Engle, N.; Bali, G.; Jawdy, S.S.; Sykes, R.W.; Davis, M.; Ragauskas, A.; Tuskan, G.A.; et al. Down-Regulation of *KORRIGAN*-Like Endo-beta-1,4-Glucanase Genes Impacts Carbon Partitioning, Mycorrhizal Colonization and Biomass Production in *Populus*. *Front. Plant Sci.* **2016**, *7*, 1455. [CrossRef]

72.　Kong, D.; Karve, R.; Willet, A.; Chen, M.K.; Oden, J.; Shpak, E.D. Regulation of plasmodesmatal permeability and stomatal patterning by the glycosyltransferase-like protein KOBITO1. *Plant Physiol.* **2012**, *159*, 156–168. [CrossRef]

73.　Foissac, S.; Sammeth, M. ASTALAVISTA: Dynamic and flexible analysis of alternative splicing events in custom gene datasets. *Nucleic Acids Res.* **2007**, *35*, W297–W299. [CrossRef]

74.　Florea, L.; Song, L.; Salzberg, S. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000Research* **2013**, *2*, 188. [CrossRef]

75.　Carvalho, R.F.; Feijao, C.V.; Duque, P. On the physiological significance of alternative splicing events in higher plants. *Protoplasma* **2013**, *250*, 639–650. [CrossRef]

76.　Chi, Y.; Yang, Y.; Zhou, Y.; Zhou, J.; Fan, B.; Yu, J.Q.; Chen, Z. Protein-protein interactions in the regulation of WRKY transcription factors. *Mol. Plant* **2013**, *6*, 287–300. [CrossRef]

77.　Li, Y.; Chen, C.; Xie, Z.; Xu, J.; Wu, B.; Wang, W. Integrated Analysis of mRNA and microRNA Elucidates the Regulation of Glycyrrhizic Acid Biosynthesis in *Glycyrrhiza uralensis* Fisch. *Int. J. Mol. Sci.* **2020**, *21*, 3101. [CrossRef]

78.　Ning, J.; Zhang, B.; Wang, N.; Zhou, Y.; Xiong, L. Increased Leaf Angle1, a Raf-Like MAPKKK That Interacts with a Nuclear Protein Family, Regulates Mechanical Tissue Formation in the Lamina Joint of Rice. *Plant Cell* **2011**, *23*, 4334–4347. [CrossRef]

79.　Huang, D.; Wang, S.; Zhang, B.; Shangguan, K.; Shi, Y.; Zhang, D.; Liu, X.; Wu, K.; Xu, Z.; Fu, X.; et al. A Gibberellin-Mediated DELLA-NAC Signaling Cascade Regulates Cellulose Synthesis in Rice. *Plant Cell* **2015**, *27*, 1681–1696. [CrossRef]

80.　Houseley, J.; Tollervey, D. The many pathways of RNA degradation. *Cell* **2009**, *136*, 763–776. [CrossRef]

81.　Wu, H.J.; Ma, Y.K.; Chen, T.; Wang, M.; Wang, X.J. PsRobot: A web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* **2012**, *40*, W22–W28. [CrossRef]

82.　Vazquez, F. *Arabidopsis* endogenous small RNAs: Highways and byways. *Trends Plant Sci.* **2006**, *11*, 460–468. [CrossRef]

83.　Yu, S.; Galvão, V.C.; Zhang, Y.C.; Horrer, D.; Zhang, T.Q.; Hao, Y.H.; Feng, Y.Q.; Wang, S.; Schmid, M.; Wang, J.W. Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. *Plant Cell* **2012**, *24*, 3320–3332. [CrossRef]

84. Damodharan, S.; Corem, S.; Gupta, S.K.; Arazi, T. Tuning of SlARF10A dosage by sly-miR160a is critical for auxin-mediated compound leaf and flower development. *Plant J.* **2018**, *96*, 855–868. [CrossRef]

85. Raman, S.; Greb, T.; Peaucelle, A.; Blein, T.; Laufs, P.; Theres, K. Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in *Arabidopsis thaliana*. *Plant J.* **2008**, *55*, 65–76. [CrossRef]

86. Bhardwaj, A.R.; Joshi, G.; Pandey, R.; Kukreja, B.; Goel, S.; Jagannath, A.; Kumar, A.; Katiyar-Agarwal, S.; Agarwal, M. A genome-wide perspective of miRNAome in response to high temperature, salinity and drought stresses in *Brassica juncea* (Czern) L. *PLoS ONE* **2014**, *9*, e92456. [CrossRef]

87. Chen, Z.H.; Bao, M.L.; Sun, Y.Z.; Yang, Y.J.; Xu, X.H.; Wang, J.H.; Han, N.; Bian, H.W.; Zhu, M.Y. Regulation of auxin response by miR393-targeted transport inhibitor response protein 1 is involved in normal development in *Arabidopsis*. *Plant Mol. Biol.* **2011**, *77*, 619–629. [CrossRef]

88. Lin, S.I.; Chiang, S.F.; Lin, W.Y.; Chen, J.W.; Tseng, C.Y.; Wu, P.C.; Chiou, T.J. Regulatory network of microRNA399 and PHO2 by systemic signaling. *Plant Physiol.* **2008**, *147*, 732–746. [CrossRef]

89. Takei, K.; Yamaya, T.; Sakakibara, H. *Arabidopsis CYP735A1* and *CYP735A2* encode cytokinin hydroxylases that catalyze the biosynthesis of *trans*-Zeatin. *J. Biol. Chem.* **2004**, *279*, 41866–41872. [CrossRef]

90. Peacock, R.B.; Hicks, C.W.; Walker, A.M.; Dewing, S.M.; Lewis, K.M.; Abboud, J.C.; Stewart, S.W.A.; Kang, C.; Watson, J.M. Structural and Functional Characterization of Dynamic Oligomerization in *Burkholderia cenocepacia* HMG-CoA Reductase. *Biochemistry* **2019**, *58*, 3960–3970. [CrossRef]

91. Hu, H.; Dong, C.; Sun, D.; Hu, Y.; Xie, J. Genome-Wide Identification and Analysis of U-Box E3 Ubiquitin-Protein Ligase Gene Family in Banana. *Int. J. Mol. Sci.* **2018**, *19*, 3874. [CrossRef] [PubMed]

92. Dong, M.Y.; Fan, X.W.; Pang, X.Y.; Li, Y.Z. Decrypting tubby-like protein gene family of multiple functions in starch root crop cassava. *AoB Plants* **2019**, *11*, plz075. [CrossRef] [PubMed]

93. Kim, M.J.; Yu, C.Y.; Theusch, E.; Naidoo, D.; Stevens, K.; Kuang, Y.L.; Schuetz, E.; Chaudhry, A.S.; Medina, M.W. SUGP1 is a novel regulator of cholesterol metabolism. *Hum. Mol. Genet.* **2016**, *25*, 3106–3116. [CrossRef] [PubMed]

94. Ran, M.; Luo, H.; Gao, H.; Tang, X.; Chen, Y.; Zeng, X.; Weng, B.; Chen, B. miR-362 knock-down promotes proliferation and inhibits apoptosis in porcine immature Sertoli cells by targeting the *RMI1* gene. *Reprod. Domest. Anim. Zuchthygiene* **2020**, *55*, 547–558. [CrossRef] [PubMed]

95. Pan, S.; Cheng, X.; Chen, H.; Castro, P.D.; Ittmann, M.M.; Hutson, A.W.; Zapata, S.K.; Sifers, R.N. ERManI is a target of miR-125b and promotes transformation phenotypes in hepatocellular carcinoma (HCC). *PLoS ONE* **2013**, *8*, e72829. [CrossRef] [PubMed]

96. Du, X.; Huang, G.; He, S.; Yang, Z.; Sun, G.; Ma, X.; Li, N.; Zhang, X.; Sun, J.; Liu, M.; et al. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat. Genet.* **2018**, *50*, 796–802. [CrossRef]

97. Maydan, G.; Noyman, I.; Har-Zahav, A.; Neriah, Z.B.; Pasmanik-Chor, M.; Yeheskel, A.; Albin-Kaplanski, A.; Maya, I.; Magal, N.; Birk, E.; et al. Multiple congenital anomalies-hypotonia-seizures syndrome is caused by a mutation in PIGN. *J. Med. Genet.* **2011**, *48*, 383. [CrossRef]

98. Li, D.; Fei, T.; Wang, Y.; Zhao, Y.; Dai, L.; Fu, X.; Li, X. A cold-active 1,4-α-glucan branching enzyme from *Bifidobacterium longum* reduces the retrogradation and enhances the slow digestibility of wheat starch. *Food Chem.* **2020**, *324*, 126855. [CrossRef]

99. Li, Y.; He, M.; Li, J.; Yao, Y.; Zhu, L.; Wu, B. Regulatory protein genes and microRNAs in response to selenium stimuli in *Pueraria lobata* (Willd.) Ohwi. *Metallomics* **2021**, *13*, mfaa004. [CrossRef]

100. Luo, X.; Reiter, M.A.; d'Espaux, L.; Wong, J.; Denby, C.M.; Lechner, A.; Zhang, Y.; Grzybowski, A.T.; Harth, S.; Lin, W.; et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **2019**, *567*, 123–126. [CrossRef]

101. Reddy, A.S.; Marquez, Y.; Kalyna, M.; Barta, A. Complexity of the alternative splicing landscape in plants. *Plant Cell* **2013**, *25*, 3657–3683. [CrossRef]

102. Liu, J.; Jung, C.; Xu, J.; Wang, H.; Deng, S.; Bernad, L.; Arenas-Huertero, C.; Chua, N.H. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*. *Plant Cell* **2012**, *24*, 4333–4345. [CrossRef]

103. Calixto, C.P.G.; Tzioutziou, N.A.; James, A.B.; Hornyik, C.; Guo, W.; Zhang, R.; Nimmo, H.G.; Brown, J.W.S. Cold-Dependent Expression and Alternative Splicing of *Arabidopsis* Long Non-coding RNAs. *Front. Plant Sci.* **2019**, *10*, 235. [CrossRef]

104. Chao, H.P.; Chen, Y.; Takata, Y.; Tomida, M.W.; Lin, K.; Kirk, J.S.; Simper, M.S.; Mikulec, C.D.; Rundhaug, J.E.; Fischer, S.M.; et al. Systematic evaluation of RNA-Seq preparation protocol performance. *BMC Genom.* **2019**, *20*, 571. [CrossRef]

105. Hafner, M.; Landgraf, P.; Ludwig, J.; Rice, A.; Ojo, T.; Lin, C.; Holoch, D.; Lim, C.; Tuschl, T. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **2008**, *44*, 3–12. [CrossRef]

106. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

107. Xu, L.N.; Ling, Y.H.; Wang, Y.Q.; Wang, Z.Y.; Hu, B.J.; Zhou, Z.Y.; Hu, F.; He, K.L. Identification of differentially expressed microRNAs between *Bacillus thuringiensis* Cry1Ab-resistant and -susceptible strains of *Ostrinia furnacalis*. *Sci. Rep.* **2015**, *5*, 15461. [CrossRef]

108. Meyers, B.C.; Axtell, M.J.; Bartel, B.; Bartel, D.P.; Baulcombe, D.; Bowman, J.L.; Cao, X.; Carrington, J.C.; Chen, X.; Green, P.J.; et al. Criteria for annotation of plant MicroRNAs. *Plant Cell* **2008**, *20*, 3186–3190. [CrossRef]

109. Zhang, B.H.; Pan, X.P.; Cox, S.B.; Cobb, G.P.; Anderson, T.A. Evidence that miRNAs are different from other RNAs. *Cell. Mol. Life Sci. CMLS* **2006**, *63*, 246–254. [CrossRef]

110. Dai, X.; Zhuang, Z.; Zhao, P.X. Computational analysis of miRNA targets in plants: Current status and challenges. *Brief. Bioinform.* **2011**, *12*, 115–121. [CrossRef]

111. German, M.A.; Luo, S.; Schroth, G.; Meyers, B.C.; Green, P.J. Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* **2009**, *4*, 356–362. [CrossRef]

112. Lottaz, C.; Iseli, C.; Jongeneel, C.V.; Bucher, P. Modeling sequencing errors by combining Hidden Markov models. *Bioinformatics* **2003**, *19* (Suppl. 2), ii103–ii112. [CrossRef]

113. Wang, L.; Park, H.J.; Dasari, S.; Wang, S.; Kocher, J.P.; Li, W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res.* **2013**, *41*, e74. [CrossRef]

114. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [CrossRef]

115. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef]

116. Li, J.; Ma, W.; Zeng, P.; Wang, J.; Geng, B.; Yang, J.; Cui, Q. LncTar: A tool for predicting the RNA targets of long noncoding RNAs. *Brief. Bioinform.* **2015**, *16*, 806–812. [CrossRef]

117. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−ΔΔC(T)) Method. *Methods* **2001**, *25*, 402–408. [CrossRef] [PubMed]