

Article

Genome-Wide Identification and Functional Prediction of Novel Drought-Responsive lncRNAs in *Pyrus betulifolia*

Jinxing Wang, Jing Lin *, Jialiang Kan, Hong Wang, Xiaogang Li, Qingsong Yang, Hui Li and Youhong Chang

Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Pomology, Jiangsu Academy of Agricultural Sciences, Nanjing 200014, China; xing521_zi@163.com (J.W.); 2014104011@njau.edu.cn (J.K.); wh811006@163.com (H.W.); xiaogangli@aliyun.com (X.L.); yng_23@163.com (Q.Y.); lihui7904@163.com (H.L.); cyh@jaas.ac.cn (Y.C.)

* Correspondence: lj84390224@126.com; Tel.: +86-25-84390224

Received: 2 May 2018; Accepted: 14 June 2018; Published: 20 June 2018



Abstract: Increasing evidence shows that long noncoding RNAs (lncRNAs) play important roles in developmental regulation and many other biological processes in plants. However, identification of lncRNAs in *Pyrus betulifolia* is limited compared with studies of functional gene expression. Using high-throughput sequencing technology, the transcriptome of *P. betulifolia* under drought stress was analyzed to identify lncRNAs. A total of 14,478 lncRNAs were identified, of which 251 were found to be drought-responsive. The putative target genes of these differentially expressed lncRNAs were significantly enriched in metabolic processes, organic substance metabolic processes, macromolecule metabolic processes, and heterocyclic compound binding. Real-time quantitative polymerase chain reaction validation suggested that the results of the RNA sequencing data analysis were reliable. This study will provide genetic resources for pear breeding and provide reference to other pomological studies.

Keywords: *Pyrus betulifolia*; drought stress; lncRNA; identification

1. Introduction

Pear (*Pyrus* spp. L) is one of the most important and popular fruit crops cultivated commercially in most temperate regions [1] and belongs to the Pomaceae subfamily of Rosaceae. The pear is a delicate fruit with a pleasant taste and smooth texture, and thus has high customer acceptance [2]. Drought is a major abiotic stress that affects pear growth, development, productivity, and quality. To improve crop performance and productivity in water-limited regions of the world, numerous studies have been conducted on drought-responsive transcriptomes and regulatory networks in several fruit trees, such as apple [3,4], cherry [5], grape [6,7], lychee [8], and pear [9].

Long noncoding RNAs (lncRNAs) are a class of RNA transcripts that do not encode proteins [10]. They are defined as more than 200 nucleotides in length with capped 5' ends and often contain spliced introns [11,12]. lncRNAs have increasingly been shown to function as important regulatory molecules, playing many critical roles in cis- and trans-regulation of genes. lncRNAs operate transcriptionally or post-transcriptionally through diverse mechanisms, including the recruitment of factors that modify chromatin or activate transcription, serving as precursors of small RNAs and affecting nuclear architecture [13]. In humans, the expression of long intergenic noncoding RNA (lincRNA) is remarkably tissue-specific compared with coding RNA, and these lincRNAs are usually coexpressed with neighboring genes [14]. During mouse postnatal testis development, there were numerous lncRNAs overlapped or were adjacent to key transcription factors or other genes related to

spermatogenesis, and most of the differentially expressed lncRNAs showed epigenetic modification marks identical with protein-coding genes [15]. lncRNAs also play vital roles in development and the stress response in plants [16–18]. A lncRNA was found to be involved in the NSR-ASCO-lncRNA regulatory module that regulates the alternative splicing patterns of several target mRNAs during root development in *Arabidopsis thaliana* [19]. In maize, 664 drought-responsive lncRNAs have been identified [20].

In order to identify novel lncRNAs, several strategies have been employed, including both computational and experimental screening methods [21,22]. With the development of next-generation sequencing, genome-wide identification of lncRNAs has been conducted for some plants, including *A. thaliana* [23], *Medicago truncatula* [24], *Gossypium* spp. [25], *Triticum aestivum* [26], *Oryza sativa* [22], and *Zea mays* [27,28]. For example, 20,163 putative lncRNAs were identified in maize using whole genome sequence annotation and RNA sequencing (RNA-seq) datasets from 30 different experiments, and more than 90% are predicted to be precursors of small RNAs. In *A. thaliana*, 6480 lincRNAs have been identified using bioinformatics [29], and in cotton, 50,566 lincRNAs were obtained by integrating high-quality RNA-seq data with high-depth stranded RNA sequencing [30].

Pears are reproduced primarily through grafting. As one of the main pear rootstocks, *Pyrus betulifolia* Bunge exhibits both drought resistance and salt tolerance [31,32] and identification of the drought resistance genes of *P. betulifolia* will be important for improving the drought resistance abilities of pears. The genome of the pear (*Pyrus bretschneideri* Rehd.) has been sequenced recently [1], but no systematic identification of lncRNAs and their responses to drought stress has been reported in *P. betulifolia*. In this study, RNA-seq was used to analyze differentially expressed lncRNAs during drought stress. Identification of drought-resistance lncRNA genes will be beneficial for pear breeding programs.

2. Materials and Methods

2.1. Plant Materials and Dehydration Treatment

Birch-leaf pear (*P. betulifolia* Bunge) seedlings were grown in the seedling beds of the Institute of Pomology at Jiangsu Academy of Agricultural Sciences, Nanjing, Jiangsu, China. Seedlings were placed in a growth chamber under a 24-h cycle: 14 h at 25 °C in the light and 10 h at 20 °C in the dark, as per our previous report [33]. Uniform and healthy six-leaf stage seedlings were inserted into a beaker containing distilled water for 2 d before dehydration treatment. Seedlings were then transferred into $\frac{1}{2}$ strength Murashige and Skoog Basal (MS) solution containing 15% polyethylene glycol (PEG) to simulate drought stress. The leaves of seedlings were collected in triplicate at 0, 24, and 48 h after treatment, rinsed with distilled water, frozen in liquid nitrogen, and stored at –80 °C until further use.

2.2. RNA Extraction, Library Construction, and lncRNA Sequencing

To construct RNA libraries, leaves were collected after 0, 24, and 48 h of PEG- treatment of six-leaf stage seedlings, using three biological replicates. Total RNA was extracted from these samples using the Total RNA Kit (Tiangen, Beijing, China) according to the manufacturer's instructions. RNA degradation and contamination was monitored using 1% agarose gels. RNA integrity was assessed using the RNA Nano 6000 Assay Kit with the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Because some lncRNAs lack a poly(A) tail, total RNA was treated to remove ribosomal RNA (rRNA) using the Epicentre Ribo-zero rRNA Removal Kit (Epicentre, Madison, WI, USA). Finally, sequencing libraries were generated using rRNA-depleted RNA with the NEBNext Ultra™ Directional RNA LibraryPrep Kit for Illumina (NEB, Ipswich, MA, USA) according to the manufacturer's recommendations. Nine strand-specific RNA libraries were submitted to the Novogene (Beijing, China) for 150 base pair (bp) paired-end sequencing on the Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA), at a depth of 120 million reads per library.

2.3. Data Analysis

2.3.1. Quality Control, Transcriptome Assembly and lncRNAs Identification

After removing reads that included adapter sequences or poly-N and low-quality sequences ($Q20 < 20$) from the raw data, clean data were obtained for further analysis. A reference genome and gene annotation files were downloaded from the pear genome website (<http://peargenome.njau.edu.cn/>). The index of the reference genome was built using Bowtie v2.0.6 [34] and clean reads were aligned to the pear genome using TopHat [35]. Based on these results, reads mapped from each sample were assembled using Cufflinks (v2.1.1) [36] with a reference-based approach.

To identify the noncoding transcripts, we used CNCI (Coding-Non-Coding-Index) [37], CPC (Coding Potential Calculator) [38], PfamScan [39] and PhyloCSF [40] software to distinguish protein coding and non-coding sequences. If a sequence was predicted to be potential coding by at least one of the tool, it was considered as a potentially coding or a pseudogene and excluded from the final dataset. Resulting sequences smaller than 200 nt were considered small noncoding RNA, while those bigger or equal than 200 nt were considered long noncoding RNA. We filtered out the low expressed transcripts and retained the transcripts characterized by FPKM (fragments per kilobase of exon per million fragments mapped) value ≥ 0.5 using Cuffquant [41].

To identify lncRNAs acting as precursors of miRNA in birch-leaf pear, the lncRNAs were aligned with the precursors of known microRNA (miRNAs) in the miRNA Base 21.0 (<http://www.mirbase.org/>) using BLAST with default parameters [42]. The lncRNAs homologous to miRNA with $>90\%$ coverage were eventually defined as miRNA precursors. Based on multiple sequence alignments, secondary structures, and covariance models, all predicted lncRNAs were classified into families that shared a common evolutionary ancestor using Rfam [43] and INFERNAL [44] with default parameters.

2.3.2. Differential Expression Analysis

The expression levels of every isoform were calculated using the FPKM method, which is calculated based on the lengths of the fragments and the read count mapped to a particular fragment. We used the Cuffdiff (v2.1.1) software [45] to calculate FPKMs of the lncRNA genes in each sample. A p -value was assigned to each gene and adjusted by the Benjamini and Hochberg approach for controlling the false discovery rate [46]. lncRNAs with an adjusted p -value < 0.05 were set as the threshold for significant differential expression [47].

2.3.3. Putative Target Gene Prediction and Enrichment Analysis

To predict the putative target genes of significantly differentially expressed lncRNAs two approaches were used, coexpression and colocation method. Coexpression were predicted based on the expression correlation coefficient (Pearson's correlation > 0.95 or < -0.95). In colocation, we searched for all the protein-coding genes 10 kb upstream or downstream of the differently expressed lncRNAs [48]. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of the putative target genes were conducted respectively. GO enrichment analysis was conducted using the Goseq R package v1.32, and GO terms with corrected p -values less than 0.05 were considered significantly enriched. KEGG orthology based annotation system (KOBAS) software was used to test for statistical significance of the differentially expressed lncRNA putative target genes in KEGG pathways (<http://www.genome.jp/kegg>) [49].

2.3.4. Quantitative Real-Time Polymerase Chain Reaction Analysis

The expression profiles of drought-responsive lncRNAs were validated through quantitative PCR. Total RNA for use as a template was extracted from leaves using the Total RNA Kit (TIANGEN) according to the manufacturer's instructions, the first cDNA strand was synthesized from 1000 ng total RNA in a volume of 20 μ L using the PrimeScript RT Reagent Kit with genomic DNA (gDNA) eraser (Perfect Real Time), (Clontech, Shiga, Japan), according to the manufacturer's protocol.

Twelve lncRNAs were randomly selected from the differentially expressed lncRNA and analyzed using Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR). Primers were designed using Primer Premier 5.0 software [50], and RT-qPCR was performed on a 7500 Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA). The total reaction volume was 20 μ L, containing 10 μ L 2 \times SYBR Premix Ex Ta (TaKaRa Bio Inc., Shiga, Japan), 1 μ L complementary DNA (cDNA) reaction mixture, 0.5 μ L of each primer with concentration is 10 pmol/ μ L, 0.5 μ L ROX Reference DyeII (50 \times), and 7.5 μ L ddH₂O. The primers sequence used in our PCR experiments are described in Supplementary Table S1 1. PCR was performed as follows: pre-denaturation at 95 °C for 30 s, denaturation at 95 °C for 3 s, annealing at 60 °C for 30 s, and 55–95 °C for melting curve analysis. All reactions were performed using biological triplicates. The $2^{-\Delta\Delta C_t}$ method was used to calculate relative changes in gene expression between control and treatment plants [51]. Ubiquitin (*UBQ*) gene was used as housekeeping gene for normalization [52].

3. Results

3.1. Identification of lncRNAs in the Birch-Leaf Pear

To investigate novel and drought-responsive lncRNAs in *P. betulifolia*, we performed Illumina sequencing of RNA from the PEG-stressed pear using HiSeq2500. In total, we obtained 877.97 million raw reads from nine samples. After filtering out adapter sequences and low-quality reads, we obtained 853.77 million clean reads for further analysis (Table 1).

Table 1. Sequencing data for nine libraries obtained by RNA sequencing.

Sample Name	Raw Reads	Clean Reads	Clean Bases	Error Rate (%)	Q20 (%)	Q30 (%)	GC Content (%)
h0-1	104,464,354	101,335,652	15.2 G	0.01	97.61	93.56	43.19
h0-2	92,734,618	90,096,636	13.51	0.02	95.63	88.03	44.27
h0-3	89,700,722	87,193,746	13.08 G	0.03	95.00	86.72	43.49
h24-1	99,690,832	96,780,130	14.52 G	0.02	97.51	93.32	43.07
h24-2	96,641,070	94,021,570	14.1 G	0.02	97.37	93.04	43.60
h24-3	99,260,444	96,754,238	14.51 G	0.02	97.57	93.46	43.59
h48-1	102,684,392	99,947,264	14.99 G	0.02	97.57	93.48	43.02
h48-2	98,612,476	95,718,804	14.36 G	0.02	97.56	93.45	43.60
h48-3	95,021,280	92,290,956	13.84 G	0.02	97.47	93.24	43.35

Q20 and Q30 are the percentages of reads with Phred scores over than 20 and 30, respectively. GC content (%) means G + C bases as the percentage of total bases. 0 h, 24 h and 48 h represented the samples that were treated under drought stress for 0 h, 24 h and 48 h respectively.

Sequence reads were mapped with TopHat [35] and assembled with Cufflinks [45]. We discovered 120,615 total transcripts. To identify pear lncRNA candidates in these 120,615 transcripts, the transcripts were further analyzed using computational and experimental methods. Five sequential stringent filter processes were used to identify lncRNAs (Figure 1). First, we filtered out transcripts with low expression and low-credibility single transcripts, and selected sequences with exon number ≥ 2 . Second, we filtered transcripts with length < 200 nucleotides. Three, using Cuffcompare software, we filtered out the transcripts that overlap with the database annotation exon region. Fourth, we filtered by expression, selecting transcripts with expression FPKM ≥ 0.5 . Fifth, we selected for coding potential, and finally obtained 14,478 reliably expressed novel lncRNAs (Supplementary File S1), including 10,896 lincRNAs and 3582 antisense lncRNAs.

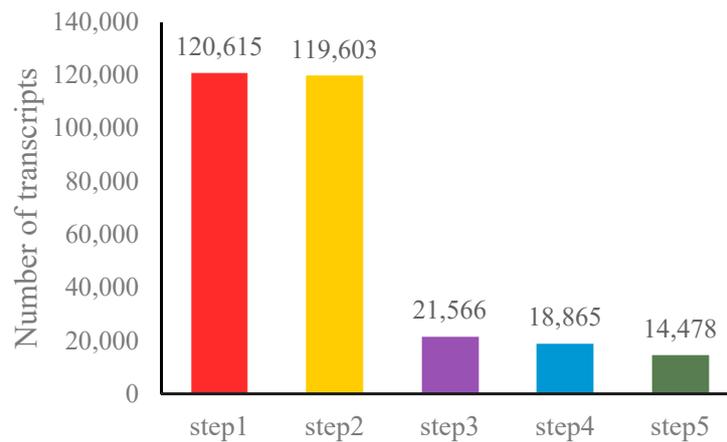


Figure 1. Statistics of candidate long noncoding RNAs (lncRNAs). Five filter processes were used to identify lncRNAs. Step1 filtered out the low expression transcripts. Step2 filtered transcripts with length <200 nucleotides. Step3 filtered out the transcripts that contain exon region. Step4 selected transcripts with FPKM (fragments per kilobase of exon per million fragments mapped) ≥ 0.5 . In Step5 we selected for coding potential, and finally 14,478 lncRNAs remained.

3.2. Characterization of Pear lncRNAs

Based on the results, 14,478 putative lncRNAs were selected and used for further analysis. The lncRNAs ranged in length from 200 to 9334 bp. The most abundant length was 200–400 bp (Figure 2). In addition, the lncRNAs distribution in pear genome was examined. There was an average of 30.4 lncRNAs per Mbp. The eighth chromosome had the highest lncRNA packing density and per Mbp nucleotides contained 41.9 lncRNAs. The lowest packing density located on the fifth chromosome and containing 17.2 lncRNAs per Mbp of nucleotides. By classify the lncRNAs into different non-coding RNA (ncRNA) families, the structure, properties and functions of these predicted lncRNAs can be better annotated and understood. All lncRNA candidate were divided into different ncRNA families using INFERNAL [44]. Global detection of the lncRNAs sequences and secondary structures indicated that most of the pear lncRNAs belong to seven conserved ncRNA families: 5SrRNA, transfer RNA (tRNA9, snoR, SNORD and mir (Supplementary Table S2 F). In total, 77 lncRNAs were detected as precursors for the 240 miRNAs, for example, five miRNA precursors (gma-MIR393c, gma-MIR393d, gma-MIR393e, gma-MIR393f and gma-MIR393g) could be aligned with lncRNA3103. The miRNA precursors are shown in Supplementary Table S3.

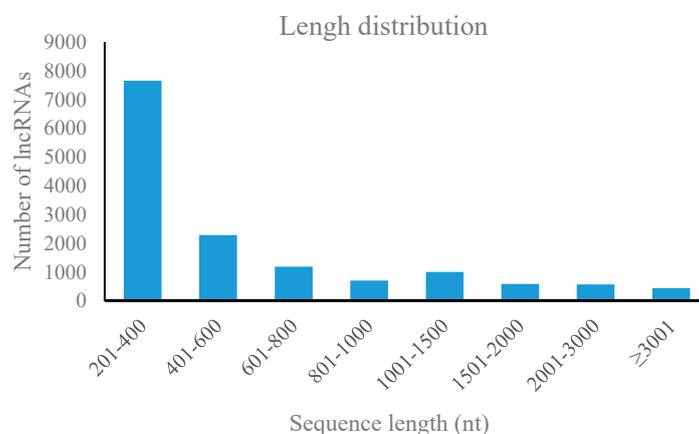


Figure 2. Length distribution of the 14,478 pear lncRNAs.

3.3. Identification of lncRNAs that Respond to Drought-Stress Treatments

To identify which of the 14,478 lncRNAs were significantly differentially expressed under drought stress treatments, the sequences were sorted using fold-change values of FPKM expression, then filtered based on false discovery rate ≤ 0.05 and $\log_2 \geq 1$. The principal component analysis (PCA) showed the reproducibility of the lncRNA expression, the h24-3 sample do not cluster with its h24-1 and h24-2 replicates (Figure 3A), so it was excluded to avoid tampering posterior expression analyses. The results indicated that the expression of 251 lncRNAs changed significantly by take the 24 h and 48 h expression values compared to 0 h (Supplementary Table S4). We identified 191 lncRNAs with significantly different expression between the 0 h and the 24 h, of which 92 genes were upregulated and 99 were downregulated. A total of 126 lncRNAs exhibited significant differential expression between 0 and 48 h, with 72 genes upregulated and 54 downregulated. The greatest number of differentially expressed transcripts was between 0 h and 24 h. The differentially expressed lncRNAs among the three stages are represented in a heat map (Figure 3B).

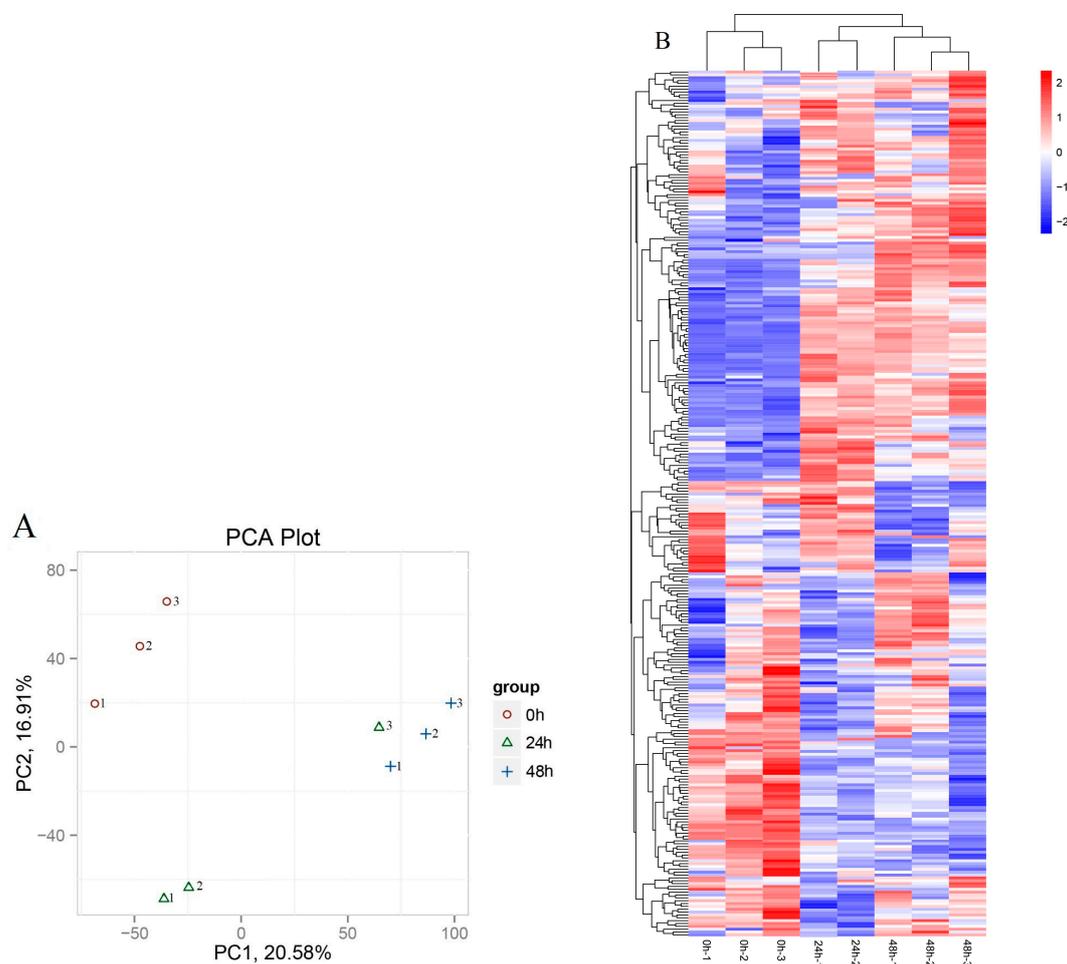


Figure 3. Principal component analysis (PCA) and heat map of lncRNAs during the different drought stages. (A) PCA of lncRNAs during different drought stage. x-axis is the contribution rate of the first principal component to the sample difference, y-axis is the contribution rate of the second principal component to the sample difference. 0 h and 48 h samples clustered together, but h24-3 does not cluster with its replicates; (B) A hierarchical clustered heat map showing the \log_{10} transformed expression values for differentially expressed lncRNAs between 0 h, 24 h and 48 h. Red represents higher expression and blue represents lower expression. 0 h, 24 h and 48 h indicate that the samples were treated under drought stress for 0 h, 24 h and 48 h, respectively.

Comparison of changes in lncRNAs between the two groups uncovered similarities and considerable differences. There are 125 unique significantly differentially expressed lncRNAs in the comparison of 24 h vs 0 h, 60 unique significantly differentially expressed lncRNAs in the comparison of 48 h vs 0 h. For the two groups, whereas 66 lncRNAs are common to both groups, indicating that lncRNAs are differentially expressed depending on the dehydration status of the plant.

3.4. Functional Prediction of Drought Response lncRNAs

To understand the potential functions of lncRNAs during drought stress, we conducted GO and KEGG enrichment analysis of the putative target genes of differentially expressed lncRNAs. We first calculated the Pearson correlation coefficient between lncRNAs and messenger RNAs (mRNAs) by examining the paired lncRNA and mRNA expression profiles. For differently expressed lncRNAs and their putative target genes between 24 h and 0 h, a total of 3139 positive and 1939 negative interaction relationships putative target genes were obtained, for 48 h and 0 h a total of 3967 positive and 1486 negative interaction relationship putative target genes were detected. GO enrichment (Figure 4A,B) analysis showed that the putative target genes of differentially expressed lncRNAs were significantly enriched in metabolic processes, organic substance metabolic processes, catalytic activity, and structural molecule activity (Supplementary Tables S5 and S6). KEGG is a database for determining the high-level functions of biological systems. KEGG pathway analysis of the putative target genes revealed that the most significant pathways were ribosome, proteasome, photosynthesis and biosynthesis of amino acids Supplementary Tables S7 and S8).

Previous studies showed that lncRNAs could play a cis-regulatory role by mediating the expression of neighboring genes. We sought the 10 kb upstream or downstream of the differentially expressed lncRNAs to find protein-coding genes. We found 3010 protein-coding genes close to 185 differentially expressed lncRNAs between 24 h and 0 h, and 2158 protein-coding genes were located close to 126 differentially expressed lncRNAs between 48 h and 0 h. GO term analysis of the putative target genes were annotated as regulating biological process, response to stimulus, and catalytic activity (Supplementary Tables S9 and S10). KEGG pathway analysis of the putative target genes showed that the most significant pathways were proteasome, photosynthesis, oxidative phosphorylation and metabolic pathways (Supplementary Tables S11 and S12).

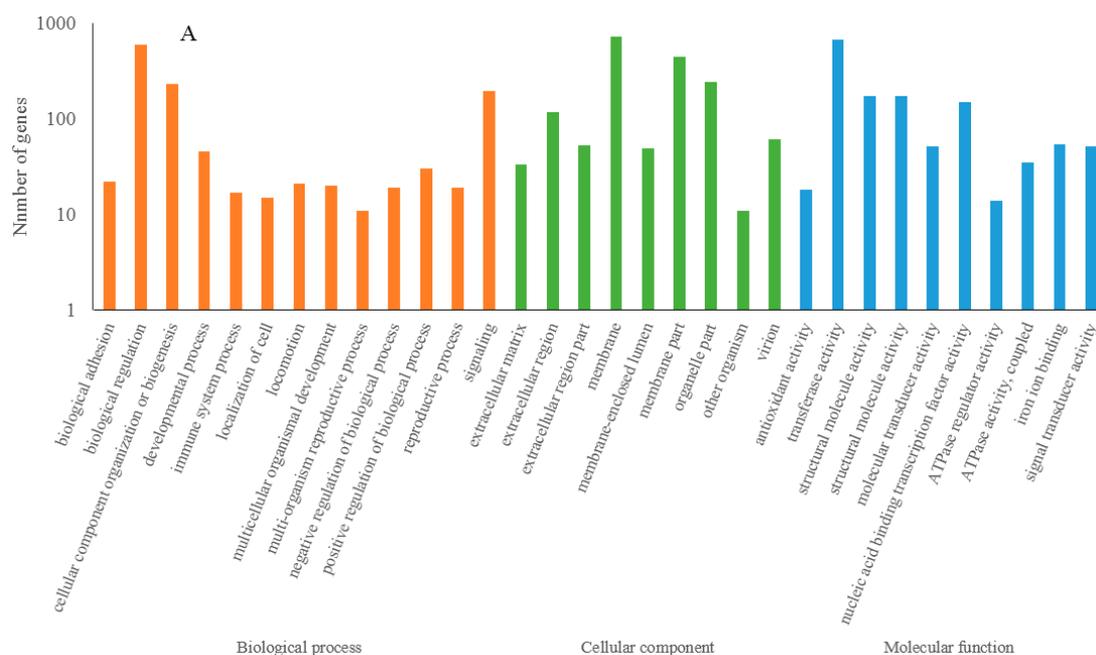


Figure 4. Cont.

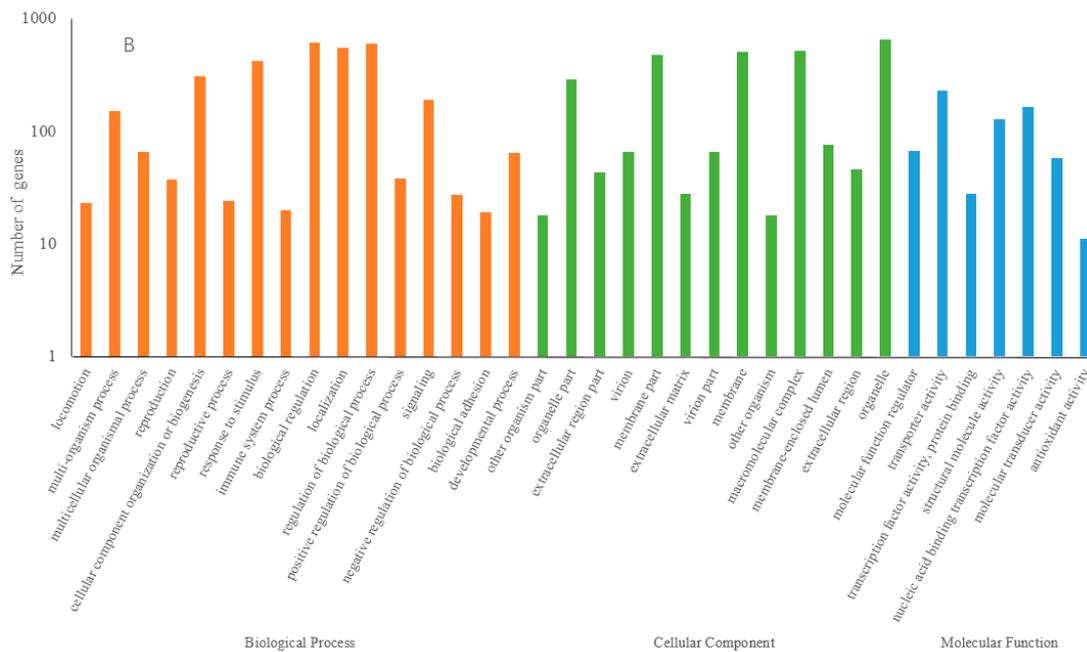


Figure 4. Functional analysis of differentially expressed lncRNA in drought stress. (A) Gene ontology (GO) analysis of putative target genes of differentially expressed lncRNA of 24 h vs. 0 h; (B) GO analysis of putative target genes of differentially expressed lncRNA of 48 h vs. 0 h. GO analysis contains three branches: biological process (orange), cellular component (green) and molecular function (blue). 0 h, 24 h and 48 h indicate that the samples were treated under drought stress for 0 h, 24 h and 48 h, respectively.

The response to drought stress in plants is a complicated process, involving several genes and metabolic network, hormones synthesis is one of important factor [53]. In total, we found that 9 lncRNAs were associated with auxin and 2 lncRNAs were associated with cytokinin. Oxidation-reduction processed pathway is another important pathway on drought stress, in this study, a total of 43 putative target genes associated with this pathway in common between the sets of co-expressed and co-located. We also identified approximately 69 putative target genes linked with ubiquitin, which may take part in signal transduction and the degradation of protein in response to stress [54] (Supplementary Tables S13 and S14).

3.5. Real-Time Quantitative Polymerase Chain Reaction Validation

To validate the reliability of the transcriptome gene expression profiles, 10 differentially expressed lncRNAs were randomly selected for expression analysis through RT-qPCR. The expression patterns shown in the RT-qPCR results were consistent with the RNA-seq results (Figure 5). For example, the relative expression of lncRNA7695 was increase on 24 h but decreased on 48 h—this was consistent with the RNA-seq result (Supplementary Table S15). We also validated the location of lncRNA7695 and lncRNA4073. They were located in scaffold1249.0 and scaffold215.0, respectively (Supplementary Figure S1). Both results suggest the reliability of the experiment data analysis in terms of expression data and sequence coverage.

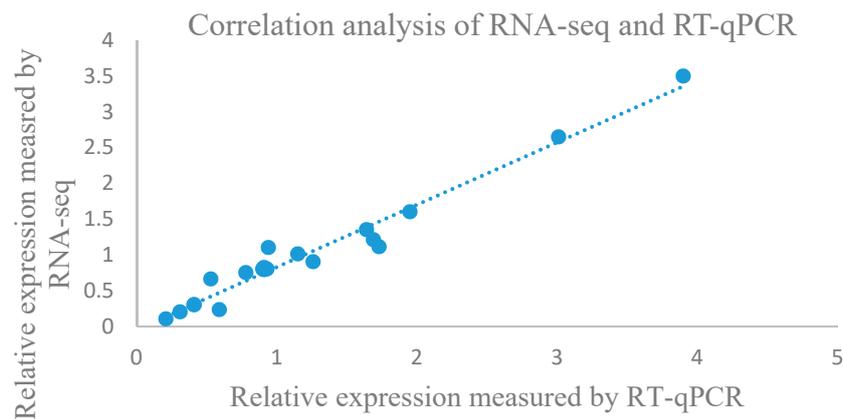


Figure 5. Confirmation of the expression patterns of lncRNAs using real-time quantitative polymerase chain reaction (RT-qPCR). Ten lncRNAs were randomly selected from differentially expressed lncRNA on different time point of drought stress. The x-axis indicates the relative expression measured by RT-qPCR. The Y-axis indicates relative expression measured by RNA sequencing (RNA-seq). The Pearson correlation of relative expression measured by RNA-seq and RT-qPCR was 0.91; ubiquitin (*UBQ*) was used as the reference gene.

4. Discussion

We obtained a large number of lncRNAs from the birch-leaf pear: 14,478 lncRNAs were reliably identified. Among other plants, 50,566, 6480, 2224, and 20,163 lncRNAs were previously identified in *Gossypium* spp. [30], *Arabidopsis* [29], rice [20], and two studies of *Zea mays* [27], respectively. The number of pear lncRNAs we identified is greater than the number identified in *Arabidopsis*, rice, but fewer than in *Gossypium* spp. and *Z. mays*. In *Gossypium* spp. and *Z. mays*, cDNA libraries were generated from multiple tissue types, whereas the cDNA libraries in the present study were generated only from leaves. Therefore, we likely identified only a subset of the total lncRNAs in birch-leaf pear, and we believe that more lncRNAs may be identified using different organs and under different stresses.

In contrast to mRNAs, lncRNAs are expressed at a low level in a tissue-specific manner, and generally exhibit poor conservation across different species [55]. lncRNAs can regulate gene expression in various ways under abiotic and biotic stresses. For example, lncRNA plays a critical role controlling the stress-responsive in *Populus* and wheat [16,56]. lncRNA173 target gene SUCROSE SYNTHASE 4 showed responsive to high temperature in *A. thaliana* [57]. In wheat, 125 lncRNAs responsive to powdery mildew infection and heat stress were identified [15]. In our result, the expression levels of 251 lncRNAs were found to change significantly during our three-stage treatment. These results indicate the drought stress damage affected the expression of some lncRNAs.

Plants are frequently affected by a wide range of environmental stressors [58]. To cope with this stress, plants have evolved physiological and biochemical mechanisms to adapt to stress [59]. To elucidate the function of putative target genes of differently expressed lncRNAs, we assigned the putative target genes of lncRNAs to various GO categories and identified several associated with the drought-stress response, such as those related to metabolic processes, organic substance metabolic processes, and catalytic activity. Using the KEGG database, we annotated the putative target genes to specific pathways, among which metabolic pathways accounted for a great proportion. In a previous study, 662 European beech genes were found to be differentially expressed under drought stress, and a large proportion of these genes were categorized into metabolic pathways [60]. Based on KEGG analysis of drought stress in Chinese fir (*Cunninghamia lanceolata*), the largest proportion of differentially expressed genes were related to metabolic pathways [61]. Thus, in this study, we obtained results consistent with those of previous research, providing further insight into the role of lncRNAs in the drought response.

In summary, we identified 14,478 lncRNAs in birch-leaf pear and showed that 251 were dehydration-responsive. These results indicate that the drought stress affected the expression of some lncRNAs. The putative target genes of differentially expressed lncRNA were identified, some of these putative target genes are involved in desiccation-response processes. The results of RT-qPCR validated the reliability of the RNA-seq. Our results provide a rich genetic resource for discovery of genes related to drought stress and can be readily applied to other fruit tree species.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/9/6/311/s1>, File S1. The 14,478 lncRNA sequences obtained by RNA sequencing of *Pyrus betulifolia* in control and drought stressed conditions. Table S1. Sequence of the primers used in qRT-PCR analysis. Table S2. The family classify of lncRNA of *Pyrus betulifolia* obtained from drought stressed condition. Table S3. lncRNA target precursors for miRNA in *Pyrus betulifolia* under drought stressed conditions. Table S4. The significantly differentially expressed lncRNAs between 24 h and 48 h compared to 0 h of *Pyrus betulifolia* under drought stress. Table S5. Gene Ontology (GO) of putative target genes of differentially expressed lncRNA of 24 h vs. 0 h based on coexpression. Table S6. Gene Ontology(GO) of putative target genes of differentially expressed lncRNA of 48 h vs. 0 h based on coexpression. Table S7. KEGG enrichment result of putative target genes of differentially expressed lncRNA of 24 h vs. 0 h based on coexpression. Table S8. KEGG enrichment result of putative target genes of differentially expressed lncRNA of 48 h vs. 0 h based on coexpression. Table S9. Gene Ontology(GO) of putative target genes of differentially expressed lncRNA of 24 h vs. 0 h based on colocation. Table S10. Gene Ontology(GO) of putative target genes of differentially expressed lncRNA of 48 h vs. 0 h based on colocation. Table S11. KEGG enrichment result of putative target genes of differentially expressed lncRNA of 24 h vs. 0 h based on colocation. Table S12. KEGG enrichment result of putative target genes of differentially expressed lncRNA of 48 h vs. 0 h based on colocation. Table S13. lncRNA and their putative target mRNA based on the coexpression results. Table S14. lncRNA and their putative target mRNA based on the colocation results. Table S15. Relative expression data measured by qRT-PCRs and RNAseq. Figure S1. The genome browser with read coverage of the lncRNA7695 and lncRNA4073.

Author Contributions: J.W. (Overall experiments and data analysis) with J.L. (Data analysis), J.K. (Data analysis), H.W. (Data analysis), X.L. (Data analysis), Q.Y. (Data analysis), H.L. (Lead principal investigator) and Y.C. (Part of paper's improvement).

Funding: This work was supported by the National Natural Sciences Foundation of China (Grant No. 31540052).

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

References

1. Wu, J.; Wang, Z.; Shi, Z.; Zhang, S.; Ming, R.; Zhu, S.; Khan, M.A.; Tao, S.; Korban, S.S.; Wang, H.; et al. The genome of the pear (*Pyrus bretschneideri* Rehd). *Genome Res.* **2013**, *23*, 396–408. [[CrossRef](#)] [[PubMed](#)]
2. Kan, J.; Liu, T.; Ma, N.; Li, H.; Li, X.; Wang, J.; Zhang, B.; Chang, Y.; Lin, J. Transcriptome analysis of Callery pear (*Pyrus calleryana*) reveals a comprehensive signalling network in response to *Alternaria alternata*. *PLoS ONE* **2017**, *12*, e0184988. [[CrossRef](#)] [[PubMed](#)]
3. Mao, K.; Dong, Q.; Li, C.; Liu, C.; Ma, F. Genome Wide Identification and Characterization of Apple bHLH Transcription Factors and Expression Analysis in Response to Drought and Salt Stress. *Front. Plant Sci.* **2017**, *8*, 480. [[CrossRef](#)] [[PubMed](#)]
4. Bassett, C.L.; Baldo, A.M.; Moore, J.T.; Jenkins, R.M.; Soffe, D.S.; Wisniewski, M.E.; Norelli, J.L.; Farrell, R.E. Genes responding to water deficit in apple (*Malus × domestica* Borkh.) roots. *BMC Plant Biol.* **2014**, *14*, 1–12. [[CrossRef](#)] [[PubMed](#)]
5. Shen, X.; Guo, X.; Zhao, D.; Zhang, Q.; Jiang, Y.; Wang, Y.; Peng, X.; Wei, Y.; Zhai, Z.; Zhao, W.; et al. Cloning and expression profiling of the PacSnRK2 and PacPP2C gene families during fruit development, ABA treatment, and dehydration stress in sweet cherry. *Plant Physiol. Biochem.* **2017**, *119*, 275–285. [[CrossRef](#)] [[PubMed](#)]
6. Cramer, G.R.; Ergül, A.; Grimplet, J.; Tillett, R.L.; Tattersall, E.A.; Bohlman, M.C.; Vincent, D.; Sonderegger, J.; Evans, J.; Osborne, C.; et al. Water and salinity stress in grapevines: Early and late changes in transcript and metabolite profiles. *Funct. Integr. Genom.* **2007**, *7*, 111–134. [[CrossRef](#)] [[PubMed](#)]
7. Moon, J.S.; Hur, Y.Y.; Jung, S.M.; Choi, Y.J.; Nam, J.C.; Park, J.G.; Koh, S.W. Transcript profiling of native Korean grapevine species *Vitis flexuosa*, exposed to dehydration and rehydration treatment. *Hortic. Environ. Biotechnol.* **2017**, *58*, 66–77. [[CrossRef](#)]

8. Shen, J.; Xiao, Q.; Qiu, H.; Chen, C.; Chen, H. Integrative effect of drought and low temperature on litchi (*Litchi chinensis* Sonn.) floral initiation revealed by dynamic genome-wide transcriptome analysis. *Sci. Rep.* **2016**, *6*, 32005. [[CrossRef](#)] [[PubMed](#)]
9. Li, K.Q.; Xu, X.Y.; Huang, X.S. Identification of Differentially Expressed Genes Related to Dehydration Resistance in a Highly Drought-Tolerant Pear, *Pyrus betulaefolia*, as through RNA-Seq. *PLoS ONE* **2016**, *11*, e0149352. [[CrossRef](#)] [[PubMed](#)]
10. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Dutttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermüller, J.; Hofacker, I.L.; et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488. [[CrossRef](#)] [[PubMed](#)]
11. Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M.F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B.W.; Cassady, J.P.; et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **2009**, *458*, 223. [[PubMed](#)]
12. Liu, J.; Wang, H.; Chua, N.H. Long noncoding RNA transcriptome of plants. *Plant Biotechnol. J.* **2015**, *13*, 319–328. [[CrossRef](#)] [[PubMed](#)]
13. Mach, J. The Long-Noncoding RNA ELENA1 Functions in Plant Immunity. *Plant Cell* **2017**, *29*, 916. [[CrossRef](#)] [[PubMed](#)]
14. Cabili, M.N.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J.L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **2011**, *25*, 1915–1927. [[CrossRef](#)] [[PubMed](#)]
15. Sun, J.; Lin, Y.; Wu, J. Long non-coding RNA expression profiling of mouse testis during postnatal development. *PLoS ONE* **2013**, *8*, e75750. [[CrossRef](#)] [[PubMed](#)]
16. Xin, M.; Wang, Y.; Yao, Y.; Song, N.; Hu, Z.; Qin, D.; Xie, C.; Peng, H.; Ni, Z.; Sun, Q. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol.* **2011**, *11*, 61. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, H.; Chen, X.; Wang, C.; Xu, Z.; Wang, Y.; Liu, X.; Kang, Z.; Ji, W. Long non-coding genes implicated in response to stripe rust pathogen stress in wheat (*Triticum aestivum*, L.). *Mol. Biol. Rep.* **2013**, *40*, 6245–6253. [[CrossRef](#)] [[PubMed](#)]
18. Jalali, S.; Bhartiya, D.; Lalwani, M.K.; Sivasubbu, S.; Scaria, V. Systematic transcriptome wide analysis of lncRNA-miRNA Interactions. *PLoS ONE* **2013**, *8*, e53823. [[CrossRef](#)] [[PubMed](#)]
19. Chen, M.; Wang, C.; Bao, H.; Chen, H.; Wang, Y. Genome-wide identification and characterization of novel lncRNAs in *Populus*, under nitrogen deficiency. *Mol. Genet. Genom.* **2016**, *291*, 1663–1680. [[CrossRef](#)] [[PubMed](#)]
20. Zhang, W.; Han, Z.; Guo, Q.; Liu, Y.; Zheng, Y.; Wu, F.; Jin, W. Identification of maize long non-coding RNAs responsive to drought stress. *PLoS ONE* **2014**, *9*, e98958. [[CrossRef](#)] [[PubMed](#)]
21. Marker, C.; Zemann, A.; Terhörst, T.; Kiefmann, M.; Kastenmayer, J.P.; Green, P.; Bachellerie, J.P.; Brosius, J.; Hüttenhofer, A. Experimental RNomics: Identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.* **2002**, *12*, 2002–2013. [[CrossRef](#)]
22. Li, L.; Wang, X.; Sasidharan, R.; Stolc, V.; Deng, W.; He, H.; Korbel, J.; Chen, X.; Tongprasit, W.; Ronald, P.; et al. Global identification and characterization of transcriptionally active regions in the rice genome. *PLoS ONE* **2007**, *2*, e294. [[CrossRef](#)] [[PubMed](#)]
23. Amor, B.B.; Wirth, S.; Merchan, F.; Laporte, P.; d’Aubenton-Carafa, Y.; Hirsch, J.; Maizel, A.; Mallory, A.; Lucas, A.; Deragon, J.M.; et al. Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res.* **2009**, *19*, 57–69. [[CrossRef](#)] [[PubMed](#)]
24. Wen, J.; Parker, B.J.; Weiller, G.F. In Silico identification and characterization of mRNA-like noncoding transcripts in *Medicago truncatula*. *Silico Biol.* **2007**, *7*, 485–505.
25. Ren, H.; Wang, G.; Chen, L.; Jiang, J.; Liu, L.; Li, N.; Zhao, J.; Sun, X.; Zhou, P. Genome-wide analysis of long non-coding RNAs at early stage of skin pigmentation in goats (*Capra hircus*). *BMC Genom.* **2016**, *17*, 67. [[CrossRef](#)] [[PubMed](#)]
26. Zhang, Y.C.; Chen, Y.Q. Long noncoding RNAs: New regulators in plant development. *Biochem. Biophys. Res. Commun.* **2013**, *436*, 111–114. [[CrossRef](#)] [[PubMed](#)]
27. Boerner, S.; McGinnis, K.M. Computational identification and functional predictions of long noncoding RNA in *Zea mays*. *PLoS ONE* **2012**, *7*, e43047. [[CrossRef](#)] [[PubMed](#)]

28. Li, L.; Eichten, S.R.; Shimizu, R.; Petsch, K.; Yeh, C.T.; Wu, W.; Chetoor, A.M.; Givan, S.A.; Cole, R.A.; Fowler, J.E.; et al. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol.* **2014**, *15*, R40. [[CrossRef](#)] [[PubMed](#)]
29. Liu, J.; Jung, C.; Xu, J.; Wang, H.; Deng, S.; Bernad, L. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in *Arabidopsis*. *Plant Cell* **2012**, *24*, 4333–4345. [[CrossRef](#)] [[PubMed](#)]
30. Wang, M.; Yuan, D.; Tu, L.; Gao, W.; He, Y.; Hu, H.; Wang, P.; Liu, N.; Lindsey, K.; Zhang, X. Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium* spp.). *New Phytol.* **2015**, *207*, 1181–1197. [[CrossRef](#)] [[PubMed](#)]
31. Zong, Y.; Sun, P.; Liu, J.; Yue, X.; Niu, Q.; Teng, Y. Chloroplast DNA-based genetic diversity and phylogeography of *Pyrus betulaefolia*, (Rosaceae) in Northern China. *Tree Genet. Genomes* **2014**, *10*, 739–749. [[CrossRef](#)]
32. Wang, H.; Lin, J.; Li, X.G.; Chang, Y. Genome-wide identification of pear HD-Zip gene family and expression patterns under stress induced by drought, salinity, and pathogen. *Acta Physiol. Plant.* **2015**, *37*, 189. [[CrossRef](#)]
33. Xu, Y.; Li, X.; Lin, J.; Wang, Z.; Yang, Q.; Chang, Y. Transcriptome sequencing and analysis of major genes involved in calcium signaling pathways in pear plants (*Pyrus calleryana* Decne.). *BMC Genom.* **2015**, *16*, 738. [[CrossRef](#)] [[PubMed](#)]
34. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
35. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*, 1105–1111. [[CrossRef](#)] [[PubMed](#)]
36. Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D.R.; Pimentel, H.; Salzberg, S.L.; Rinn, J.L.; Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols.* **2012**, *7*, 562. [[CrossRef](#)] [[PubMed](#)]
37. Sun, L.; Luo, H.; Bu, D.; Zhao, G.; Yu, K.; Zhang, C.; Liu, Y.; Chen, R.; Zhao, Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* **2013**, *41*, e166. [[CrossRef](#)] [[PubMed](#)]
38. Kong, L.; Zhang, Y.; Ye, Z.Q.; Liu, X.Q.; Zhao, S.Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345. [[CrossRef](#)] [[PubMed](#)]
39. Finn, R.D.; Mistry, J.; Tate, J.; Coggill, P.; Heger, A.; Pollington, J.E.; Gavin, O.L.; Gunasekaran, P.; Ceric, G.; Forslund, K.; et al. The Pfam protein families database. *Nucleic Acids Res.* **2012**, *40*, D290–D301.
40. Lin, M.F.; Jungreis, I.; Kellis, M. PhyloCSF: A comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **2011**, *27*, i275–i282. [[CrossRef](#)] [[PubMed](#)]
41. Wang, Q.; Zhang, Y.; Guo, W.; Liu, Y.; Wei, H.; Yang, S. Transcription analysis of cochlear development in minipigs. *Acta Oto-Laryngol.* **2017**, *1*. [[CrossRef](#)]
42. Lobo, I. Basic Local Alignment Search Tool (BLAST). *J. Mol. Biol.* **2008**, *215*, 403–410.
43. Bateman, A.; Birney, E.; Cerruti, L.; Durbin, R.; Griffiths-Jones, S.; Sonnhammer, E.L. The Pfam protein families database. *Nucleic Acids Res.* **2002**, *28*, 263–266. [[CrossRef](#)]
44. Nawrocki, E.P. Annotating Functional RNAs in Genomes Using Infernal. In *RNA Sequence, Structure, and Function: Computational and Bioinformatic Methods*; Humana Press: New York, NY, USA, 2014; pp. 163–197.
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](#)] [[PubMed](#)]
46. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple hypothesis testing. *J. R. Stat. Soc. B* **1995**, *57*, 289–300.
47. Deng, F.; Zhang, X.; Wang, W.; Yuan, R.; Shen, F. Identification of *Gossypium hirsutum* long non-coding RNAs (lncRNAs) under salt stress. *BMC Plant Biol.* **2018**, *18*, 23. [[CrossRef](#)] [[PubMed](#)]
48. Zhu, B.; Xu, M.; Shi, H.; Gao, X.; Liang, P. Genome-wide identification of lncRNAs associated with chlorantraniliprole resistance in diamondback moth *Plutella xylostella* (L.). *BMC Genom.* **2017**, *18*, 380. [[CrossRef](#)] [[PubMed](#)]
49. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2008**, *36*, D480–D484. [[CrossRef](#)] [[PubMed](#)]

50. Wang, J.; Abbas, M.; Wen, Y.; Niu, D.; Wang, L.; Sun, Y.; Li, Y. Selection and validation of reference genes for quantitative gene expression analyses in black locust (*Robinia pseudoacacia* L.) using real-time quantitative PCR. *PLoS ONE* **2018**, *13*, e0193076. [[CrossRef](#)] [[PubMed](#)]
51. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Method* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
52. Xu, Y.; Li, H.; Li, X.; Lin, J.; Wang, Z.; Yang, Q.; Chang, Y. Systematic selection and validation of appropriate reference genes for gene expression studies by quantitative real-time PCR in pear. *Acta Physiol. Plant.* **2015**, *37*, 40. [[CrossRef](#)]
53. Lu, X.; Wang, X.; Chen, X.; Shu, N.; Wang, J.; Wang, D.; Wang, S.; Fan, W.; Guo, L.; Guo, X.; et al. Single-base resolution methylomes of upland cotton (*Gossypium hirsutum* L.) reveal epigenome modifications in response to drought stress. *BMC Genom.* **2017**, *18*, 297. [[CrossRef](#)] [[PubMed](#)]
54. Lu, X.; Chen, X.; Mu, M.; Wang, J.; Wang, X.; Wang, D.; Yin, Z.; Fan, W.; Wang, S.; Guo, L.; et al. Genome-wide analysis of long noncoding RNAs and their responses to drought stress in cotton (*Gossypium hirsutum* L.). *PLoS ONE* **2016**, *11*, e0156723. [[CrossRef](#)] [[PubMed](#)]
55. Chekanova, J.A. Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* **2015**, *27*, 207–216. [[CrossRef](#)] [[PubMed](#)]
56. Shuai, P.; Liang, D.; Tang, S.; Zhang, Z.; Ye, C.Y.; Su, Y.; Xia, X.; Yin, W. Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. *J. Exp. Bot.* **2014**, *65*, 4975–4983. [[CrossRef](#)] [[PubMed](#)]
57. Di, C.; Yuan, J.; Wu, Y.; Li, J.; Lin, H.; Hu, L.; Zhang, T.; Qi, Y.; Gerstein, M.B.; Guo, Y.; et al. Characterization of stress-responsive lincRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. *Plant J.* **2014**, *80*, 848–861. [[CrossRef](#)] [[PubMed](#)]
58. Tang, J.; Lin, J.; Li, H.; Li, X.; Yang, Q.; Cheng, Z.M.; Chang, Y. Characterization of CIPK Family in Asian Pear (*Pyrus bretschneideri* Rehd) and Co-expression Analysis Related to Salt and Osmotic Stress Responses. *Front. Plant Sci.* **2016**, *7*, 1361. [[CrossRef](#)] [[PubMed](#)]
59. Pastori, G.M.; Foyer, C.H. Common components, networks, and pathways of cross-tolerance to stress. The central role of “redox” and abscisic acid-mediated controls. *Plant Physiol.* **2002**, *129*, 460–468. [[CrossRef](#)] [[PubMed](#)]
60. Müller, M.; Seifert, S.; Lübke, T.; Leuschner, C.; Finkeldey, R. De novo transcriptome assembly and analysis of differential gene expression in response to drought in European beech. *PLoS ONE* **2017**, *12*, e0184167. [[CrossRef](#)] [[PubMed](#)]
61. Hu, R.; Wu, B.; Zheng, H.; Hu, D.; Wang, X.; Duan, H.; Sun, Y.; Wang, J.; Zhang, Y.; Li, Y. Global Reprogramming of Transcription in Chinese Fir (*Cunninghamia lanceolata*) during Progressive Drought Stress and after Rewatering. *Int. J. Mol. Sci.* **2015**, *16*, 15194–15219. [[CrossRef](#)] [[PubMed](#)]

