Alternative Splicing Enhances the Transcriptome Complexity of *Liriodendron chinense*

Zhonghua Tu[1,2], Yufang Shen[1,2], Shaoying Wen[1,2], Yaxian Zong[1,2] and Huogen Li[1,2]*

[1] Key Laboratory of Forest Genetics & Biotechnology of Ministry of Education, Nanjing Forestry University, Nanjing, China,
[2] Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

Alternative splicing (AS) plays pivotal roles in regulating plant growth and development, flowering, biological rhythms, signal transduction, and stress responses. However, no studies on AS have been performed in *Liriodendron chinense*, a deciduous tree species that has high economic and ecological value. In this study, we used multiple tools and algorithms to analyze transcriptome data derived from seven tissues *via* hybrid sequencing. Although only 17.56% (8,503/48,408) of genes in *L. chinense* were alternatively spliced, these AS genes occurred in 37,844 AS events. Among these events, intron retention was the most frequent AS event, producing 1,656 PTC-containing and 3,310 non-PTC-containing transcripts. Moreover, 183 long noncoding RNAs (lncRNAs) also underwent AS events. Furthermore, weighted gene coexpression network analysis (WGCNA) revealed that there were great differences in the activities of transcription and post-transcriptional regulation between pistils and leaves, and AS had an impact on many physiological and biochemical processes in *L. chinense*, such as photosynthesis, sphingolipid metabolism, fatty acid biosynthesis and metabolism. Moreover, our analysis showed that the features of genes may affect AS, as AS genes and non-AS genes had differences in the exon/intron length, transcript length, and number of exons/introns. In addition, the structure of AS genes may impact the frequencies and types of AS because AS genes with more exons or introns tended to exhibit more AS events, and shorter introns tended to be retained, whereas shorter exons tended to be skipped. Furthermore, eight AS genes were verified, and the results were consistent with our analysis. Overall, this study reveals that AS and gene interaction are mutual—on one hand, AS can affect gene expression and translation, while on the other hand, the structural characteristics of the gene can also affect AS. This work is the first to comprehensively report on AS in *L. chinense*, and it can provide a reference for further research on AS in *L. chinense*.

Keywords: *L. chinense*, hybrid sequencing, alternative splicing, long noncoding RNAs, serine/arginine-rich genes, heterogeneous nuclear ribonucleoprotein genes, premature termination codon, weighted gene coexpression network analysis

INTRODUCTION

Alternative splicing (AS) is an important post-transcriptional regulation mechanism that can result in the production of multiple mature mRNAs from a single gene, and it substantially increases the plasticity of the transcriptome and the diversity of the proteome (Marquez et al., 2012; Vitulo et al., 2014). Moreover, AS can influence the stability of mRNA by introducing a premature termination codon (PTC) into the coding sequence; transcripts with PTCs are degraded through the nonsense-mediated decay (NMD) pathway (Wang et al., 2014; Szakonyi and Duque, 2018; Chaudhary et al., 2019a). Furthermore, AS plays roles in regulating gene expression by changing the efficiency of transcription elongation and/or translation (Laloum et al., 2018; Szakonyi and Duque, 2018). Although the functions of alternative splicing in animals and plants are similar, some differences exist between them in some aspects. For example, intron retention (IR) has been reported to be the most frequent AS event in plants, whereas exon skipping (ES) is the major type in humans and animals (Xu et al., 2002; Wang et al., 2008; Ruan et al., 2018; Ma et al., 2019; Wang Z. et al., 2019). The *Suppressor of Morphological defects on Genitalia* (*SMG*) gene family (*SMG 1–7*), which plays an important role in the AS-coupled-NMD pathway, has been widely reported in animals; however, no *SMG 5* and *SMG 6* paralogs have been identified in plants. Even in plants, *SMG 1* does not appear in *Arabidopsis thaliana* but rather in *Physcomitrella patens* (Riehs et al., 2008; Lloyd and Davies, 2013; Capitao et al., 2018). These studies indicate that AS has both universality and species diversity.

In plants, AS is a universal phenomenon with pivotal roles in regulating plant growth and development, flowering, biological rhythms, signal transduction, and stress responses (Filichkin et al., 2015; Sun and Xiao, 2015; Gallegos, 2018; Ruan et al., 2018; Szakonyi and Duque, 2018; Wang et al., 2018; Calixto et al., 2019; Park et al., 2019; Yang et al., 2019). For example, in peanut (*Arachis hypogaea*), *Arabidopsis thaliana*, soybean (*Glycine max*), tomato (*Solanum lycopersicum*) and moso bamboo (*Phyllostachys edulis*), approximately 37%, 61%, 63%, 65%, and 49% of multi-exon genes undergo AS events, respectively (Marquez et al., 2012; Shen et al., 2014; Ruan et al., 2018; Zhao et al., 2018; Clark et al., 2019). Furthermore, researchers have found that AS also plays important roles in male and female gametogenesis, seed germination in *A. thaliana*, plant-pathogen interactions in wheat (*Triticum aestivum*), and mineral nutrient homeostasis maintenance in rice (*Oryza sativa*) (Dong et al., 2018; Forsthoefel et al., 2018; Tognacca et al., 2019; Zhang H. et al., 2019). These findings indicate that AS is essential for plant survival and reproduction. However, no systematic studies on AS in *L. chinense* have been conducted.

Splicing factors (SFs), such as serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoprotein (hnRNP), play pivotal roles in regulating AS (Chen and Manley, 2009). SR proteins are RNA-binding proteins containing one or two motifs (RNA-recognition motif) and one C-terminal domain (arginine- and serine-rich domain) (Mai et al., 2016). hnRNPs usually contain RNA-binding domains at the N-terminal and a glycine-rich domain at the C-terminal (Talukdar et al., 2011). SR proteins regulate AS by participating in spliceosome formation, whereas hnRNPs function as activators or repressors (Mauger et al., 2008; Mai et al., 2016). However, the influence of SR proteins and hnRNPs on AS genes in *L. chinense* has not been illustrated.

Long noncoding RNAs (lncRNAs) are RNA molecules with lengths greater than 200 nt that lack the ability to encode proteins (Liu et al., 2018). In plants, lncRNAs are critical regulators of growth and development, the stress response, flowering, and other biological processes (Karlik and Gozukirmizi, 2018; Liu et al., 2018; Ayachit et al., 2019; Wang P. et al., 2019). For instance, researchers have found that *lncWOX5* and *lncWOX11* participate in the regulation of poplar root development, and *lncWOX5* negatively regulates *WOX5* while *lncWOX11* positively regulates *WOX11* (Liu et al., 2018). In barley (*Hordeum vulgare*), the lncRNA *AK370841* participates in the vitamin B6 salvage pathway (Karlik and Gozukirmizi, 2018). Moreover, lncRNAs can undergo AS. Researchers have found that during rice seed development, AS events occur in lncRNAs with complex gene structures, and exon retention (as opposed to ES) is enhanced in embryos (Kiegle et al., 2018). A study on *A. thaliana* revealed that AS of some lncRNAs is closely related to small temperature changes (Calixto et al., 2019). However, research on AS of lncRNAs in *Liriodendron chinense* has not been reported, and this knowledge gap needs to be bridged.

With the development of sequencing technologies, an increasing number of studies have been conducted on the transcriptomes of different species. Such species include lotus (*Nelumbo nucifera*), tomato, peanut, maize (*Zea mays*), moso bamboo, *Passiflora edulis* Sims, and *Cinnamomum camphora* (Liu et al., 2017; Min, 2017; Chen C. et al., 2018; Ruan et al., 2018; Wang et al., 2018; Zhao et al., 2018; Clark et al., 2019; Zhang Y. et al., 2019). Moreover, using weighted gene coexpression network analysis (WGCNA) to reveal the relationship between genes based on sequencing data has become very popular. WGCNA is a systematic biology method applied in constructing gene coexpression networks which can be further divided into different coexpression modules (Zhang and Horvath, 2005; Langfelder and Horvath, 2008). Genes from the same coexpression module usually have similar expression patterns and closely related functions (Jia et al., 2020). Researchers can associate coexpression modules with different traits to reveal the relationships between gene sets and phenotypes, and this method had been widely applied in the fields of medicine and plant science (Wang J.-L. et al., 2019; Zhu et al., 2019; Esmaeili et al., 2020; Jia et al., 2020). However, using WGCNA to reveal the AS gene coexpression module in *L. chinense* had not been reported.

L. chinense, a deciduous tree species with beautiful flowers is widely distributed in northern Vietnam and southern China (Hao et al., 1995). Due to its high economic and ecological value, many studies have been conducted on *L. chinense* (Xu et al., 2017; Tu et al., 2019). Transcriptome analyses of *L. chinense* have also been performed (Yang et al., 2014; Workman et al., 2018). Importantly, sequencing of the *L. chinense* genome was completed in 2018 (Chen J. et al., 2018). Although researchers have obtained large amounts of transcriptome data and genome data, none have studied AS in *L. chinense*. The numbers of genes

and lncRNAs that undergo AS, the frequencies of different AS events, and potential differences in AS between different tissues in *L. chinense* remain unknown.

To further investigate these topics, we used Illumina sequencing combined with single-molecule long-read sequencing technology (single-molecule real-time [SMRT] sequencing) to obtain transcriptome data for seven *L. chinense* tissues. After processing the data with multiple software programs and tools, we investigated the AS events that occurred in genes and lncRNAs, identified the differences between AS genes and non-AS genes in *L. chinense*, and determined the factors that may have caused these differences. Notably, this is the first study on AS in *L. chinense*. Our findings provide new insights into the transcriptome complexity of *L. chinense* and will serve as a reference for subsequent studies of AS in *L. chinense*.

MATERIALS AND METHODS

Plant Materials and RNA Isolation

An adult *L. chinense* tree (26 years old) from a provenance trial plantation in Xiashu, Jurong County, Jiangsu Province (119°13′ E, 32°7′N), Lushan Mountain, was used as a source tree, and Jiangxi Province (116°0′E, 29°32′N) was the provenance of the source tree. We collected seven *L. chinense* tissue type: leaves, shoot apices, sepals, bracts, petals, pistils and stamens. Each type of sample had three replicates. All 21 samples were stored in a −80°C freezer after being flash-frozen in liquid nitrogen. An RNAprep Pure Plant Kit (Tiangen, Beijing, China) was used to isolate total RNA under the direction of the product manual. Then, we used 1% agarose gels to detect the degradation and contamination of RNA, and RNA quality was detected on a NanoPhotometer® spectrophotometer (IMPLEN, CA, USA). The Qubit®2.0 Fluorometer (Life Technologies, CA, USA) was applied to measure the RNA concentration, and the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) was applied to evaluate the integrity of RNA. Only RNA samples with high quality and integrity (integrity number > 0.75) were used for library construction.

Construction of Libraries and Transcriptome Sequencing

For SMRT sequencing, total RNA from 21 samples was mixed equally. Then, a SMARTer™ PCR cDNA Synthesis Kit (Clontech, CA, USA) was used to synthesize cDNA from three micrograms of total RNA. Subsequently, we performed polymerase chain reaction (PCR) amplification, followed by fragment selection to classify the fragments into two categories, ≤ 4 kb and > 4 kb, by using the BluePippin Size Selection System (Sage Science, MA, USA). After damage repair, end repair, linker ligation, exonuclease digestion, primer binding, and DNA polymerase binding, we constructed two SMRT bell libraries. When the construction of SMRT libraries were completed, a Qubit®2.0 Fluorometer (Life Technologies, CA, USA) and an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) were used to quantify and measure the libraries, respectively.

Then, the two libraries were used to perform SMRT sequencing on a PacBio Sequel platform (Pacific Bioscience, CA, USA).

For Illumina sequencing, a NEBNext® Ultra™ RNA Library Prep Kit (NEB, Frankfurt, Germany) was used to construct libraries from three micrograms of total RNA. Briefly, we used magnetic beads with Oligo (dT) to enrich mRNA. After that, fragmentation buffer was used to break the mRNA into short fragments. Then, these mRNA fragments were used as a template with random hexamers to synthesize first-strand cDNA. After adding buffer, dNTPs, and DNA polymerase I, the second-strand cDNA was synthesized, and the AMPure XP system (Beckman Coulter, CA, USA) was used to purify double-stranded cDNA. After undergoing terminal repair, NEBNext Adaptor ligation, and size selection, the cDNA fragments were used as input materials for PCR amplification followed by PCR product purification. When the construction of the library was completed, the Bioanalyzer 2100 system (Agilent Technologies, CA, USA) was used to evaluate the quality of the library. We constructed a total of 21 Illumina libraries. Then, Illumina sequencing was performed on the Illumina HiSeq 2500 platform (Illumina Inc., CA, USA).

Hybrid Sequencing Data Analysis

The raw SMRT sequencing data were processed with SMRT Link 6.0 software (https://www.pacb.com/support/software-downloads) to select suitable subreads with lengths ≥ 50 bp and read scores ≥ 0.65. Then, the suitable subreads were used to select circular consensus sequences (CCSs) with pass values ≥ 1 and predicted accuracy values ≥ 0.8. CCSs with a 5′ primer, 3′ primer, and poly A tail were classified as a full-length read. Consensus reads were then obtained using iterative clustering for error correction (ICE), and Quiver was applied to further polish the consensus reads that could produce high-quality consensus reads (length > 200, high-quality > 0.99) (Eid et al., 2009). Finally, high-quality and nonredundant isoforms were obtained by using CD-HIT to filter redundant isoforms (Li and Godzik, 2006).

The raw reads from Illumina sequencing were processed using in-house Perl scripts to obtain clean data (PRJNA559687). High-quality clean reads did not contain adaptors or poly(-N) sequences. Then, the nucleotide errors of the SMRT high-quality consensus reads were corrected using the clean Illumina data with LoRDEC (http://atgc.lirmm.fr/lordec) software (parameter: -k 23; -s 3) (Salmela and Rivals, 2014).

Reads Map to Reference Genome and Gene Function Annotation

In this study, we used GMAP (http://research-pub.gene.com/gmap/) software to map the corrected consensus reads to the reference genome with the following parameters: no-chimeras; -n 1 (Wu and Watanabe, 2005). A portion of the corrected consensus reads were mapped in the un-annotated region in the reference genome, and we defined these reads as novel genes. Moreover, to identify the functions of all genes, the Gene Ontology (GO) database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database, EuKaryotic Orthologous Groups

(KOG) database, nucleotide (NT) database, nonredundant (NR) database, Protein Family (Pfam) database, and SWISS-PROT database, were used to annotate all genes (Kanehisa and Goto, 1999; Tatusov et al., 2003; Pruitt et al., 2007; Young et al., 2010; Finn et al., 2016; The UniProt, 2017). The NR, KOG, SWISS-PROT, and KEGG database annotation was performed on Diamond blastx (https://github.com/bbuchfink/diamond) software (parameters: more-sensitive –k 10; -e 1e-5; -f 6; -p 4) (Buchfink et al., 2015); NT database annotation was performed on blastn (ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) software (parameters: -outfmt 6; -evalue 1e-5; -max_target_seqs 10; -num_threads 4) (Altschul et al., 1997); and Pfam database annotation was performed on Hmmscan (http://hmmer.org/download.html) software (parameters: –acc; –domtblout) (Finn et al., 2011).

Identification of Differentially Expressed Genes

Cuffdiff (http://cole-trapnell-lab.github.io/cufflinks/) software was used to calculate the fragments per kilobase of transcript sequence per million mapped reads (FPKM) values of all transcripts (parameters: FDR 0.05; library-norm-method geometric; dispersion-method pooled) (Sreya and Chon-Kit Kenneth, 2016). Based on the FPKM values, we used DESeq2 (http://www.bioconductor.org/packages/release/bioc/html/DESeq.html) software to identify the differentially expressed genes with adjusted P-values smaller than 0.05.

Identification of AS Events and GO/KEGG Enrichment Analyses of AS Genes

In this study, AS events were identified using SUPPA (https://bitbucket.org/regulatorygenomic-supf/suppa) software with the default parameters (Abdel-Ghany et al., 2016). The SUPPA software classified basic AS events into seven types: IR, mutual exon exclusion (MX), alternative 5′ splicing (A5), alternative 3′ splicing (A3), alternative first exon splicing (AF), alternative last exon splicing (AL), and ES (**Supplementary Material 1**). To better understand the functions of AS genes and relationships with their products, we performed GO/KEGG enrichment analyses using GOseq software (http://www.bioconductor.org/packages/release/bioc/html/goseq.html) and KOBAS software (http://kobas.cbi.pku.edu.cn/download.php). GO terms and KEGG terms with corrected P-values less than 0.05 were considered significantly enriched terms.

Identification of Alternatively Spliced lncRNAs

In this study, we used the Coding Potential Calculator (CPC, http://cpc.cbi.pku.edu.cn/), the Coding-Non-Coding Index (CNCI, https://github.com/www-bioinfo-org/CNCI), the Pfam database, and the Predictor of Long noncoding RNAs and mEssenger RNAs based on an improved K-mer scheme (PLEK, https://sourceforge.net/projects/plek/) to identify lncRNAs with the default parameters (Kong et al., 2007; Sun et al., 2013; Li. et al., 2014; Finn et al., 2016). Transcripts were identified as putative lncRNAs only if they were predicted to be

lncRNAs by all four tools. Then we used SUPPA software to identify lncRNAs that underwent AS events.

WGCNA Analysis and Hub Gene Identification

A WGCNA package in R software was used to construct the AS gene coexpression network. Only genes ranked in the top 60% based on the ranking of the variance of FPKM values could be used for the WGCNA analysis. The pickSoftThreshold function was applied to calculate the soft threshold power, and we selected the most suitable power (16 in this study) for gene coexpression network construction. The gene coexpression network was constructed using the blockwiseModules function with the following parameters: the power was 16, the TOM type was signed, and the min module size was 300, and the other parameters were the default values.

For hub gene identification, the degree algorithm in the CytoHubba package of Cytoscape software (http://www.cytoscape.org/download.html) was used to identify hub genes in the six modules based on the WGCNA analysis results. In each module, only the top 10% genes in the degree algorithm were considered as hub genes. Hub genes and their related gene were selected for network visualization using Gephi software (https://gephi.org/).Validation of AS events using reverse transcription PCR

To validate whether AS events truly occurred, we randomly selected 8 AS genes and designed specific PCR primers using Oligo 7 software (https://en.freedownloadmanager.org/Wind-owsPC/OLIGO.html). All primers are shown in **Supplementary Material 2**. After synthesizing cDNA using the PrimeScript™ RT Master Mix (TaKaRa, Dalian, China), 20 μl of cDNA was diluted 10 times. Then, 2 μl of diluted cDNA was used as the template for PCR according to the instructions of a Phanta® Max Super-Fidelity DNA Polymerase Kit (Vazyme, Nanjing, China). The PCR procedure was as follows: 95°C pre-denaturation for 3 min; 30 cycles of denaturation (95°C for 15 s), annealing (57.5–62.5°C for 15 s), and elongation (72°C, elongation rate of 1 kb per min); and 72°C complete elongation for 5 min. Finally, 1.5 percent (w/v) agarose gel electrophoresis was used to verify the amplification result.

RESULTS

Overview of the Transcriptome and AS in *L. chinense*

Using hybrid sequencing, we detected a total of 48,408 genes, including 13,139 novel genes (**Figure 1A**). Among these genes, 13,353 had a single exon while 35,055 had multiple exons (**Figure 1A**). Moreover, upon conducting functional annotation with seven databases, we found that the NR database annotated the highest number of genes (38,087), followed by the KEGG (37,016), GO (28,690), Pfam (28,690), SWISS-PROT (26,437), NT (23,887) and KOG (20,679) databases (**Figure 1A**).

Through analysis with the SUPPA tool, we found that only 8,503 genes (210 single-exon genes and 8,293 multi-exon genes) underwent AS, and most AS genes (8,389) could be found
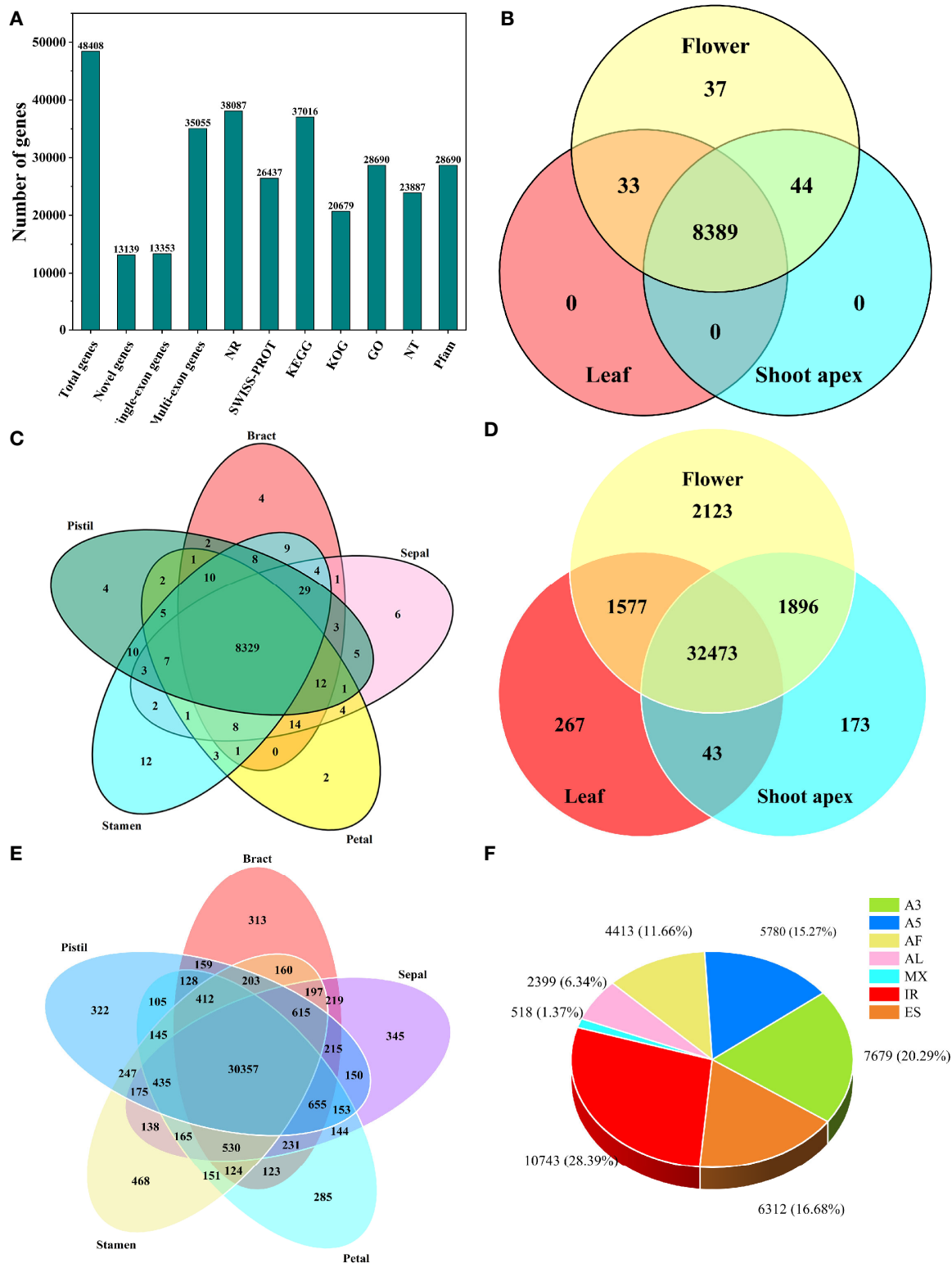
FIGURE 1 | Information about the transcriptome and alternative splicing (AS) in *L. chinense*. **(A)** Statistics for genes and their annotation in seven databases. **(B)** Distribution of alternatively spliced genes in flower, leaf and shoot apex tissues. **(C)** Distribution of alternatively spliced genes in pistil, stamen, petal, sepal and bract tissues. **(D)** Distribution of alternatively spliced transcripts in flower, leaf and shoot apex tissues. **(E)** Distribution of alternatively spliced transcripts in pistil, stamen, petal, sepal and bract tissues. **(F)** The proportions of seven AS events in *L. chinense*.

simultaneously in leaf, flower and shoot apex tissues while 37 AS genes were found only in flowers (**Figure 1B**). In pistils, stamens, petal, sepals and bracts, we found 4, 12, 2, 6, and 4 specifically expressed AS genes, respectively (**Figure 1C**). Intriguingly, we found that the number of tissue-specific AS transcripts was greater than the number of tissue-specific AS genes. Among the 2,123 tissue-specific AS transcripts, only 267 and 173 tissue-specific AS transcripts were detected in leaves and shoot apices, respectively (**Figure 1D**). Focusing on flower tissues, stamens had the largest number of tissue-specific AS variants (468), followed by sepals (345), pistils (322), bracts (313) and petals (285) (**Figure 1E**). Moreover, we identified 37,844 AS events from these 8,503 AS genes (**Figure 1F**). The most common AS event was IR (10,743, 28.39%), followed by A3 (7,679, 20.29%), ES (6,312, 16.68%), A5 (5,780, 15.27%), AF (4,413, 11.66%), and AL (2,399, 6.34%), while MX (518, 1.37%) was the least common AS event (**Figure 1F**). In other plants, such as peanut, tomato, and moso bamboo, IR has been found to be the predominant AS event (Sun and Xiao, 2015; Ruan et al., 2018; Zhao et al., 2018). Thus, IR may have a great influence in the transcriptome complexity of *L. chinense*.

GO and KEGG Enrichment Analyses of AS Genes

To further understand the functions of AS genes, we conducted GO and KEGG enrichment analyses of 8,503 AS genes. The GO enrichment analysis revealed that the top six terms were "catalytic activity", "primary metabolic process", "single-organism process", "single-organism cellular process", "cell", and "cell part" (**Figure 2A**), indicating that these AS genes play roles in basic physiological and biochemical processes. Moreover, the KEGG pathway enrichment analysis revealed that these AS genes participated in 121 pathways, most of which are associated with metabolism, such as carbon metabolism, purine metabolism, and pyrimidine metabolism (**Figure 2B**). In plants, AS plays important roles in regulating growth and development, flowering, the stress response, and other biological processes (Filichkin et al., 2015; Szakonyi and Duque, 2018). Therefore, unsurprisingly, the AS genes participated in multiple pathways and were associated with multiple biological functions.

Expression Levels of AS Genes and Splice Variants

To better understand the expression levels of AS genes and splice variants in different tissues, we performed a hierarchical clustering analysis. We found that most AS genes were strongly expressed in pistil and shoot apex tissues but weakly expressed in leaf tissues (**Figure 3A**). However, many AS genes were strongly expressed in leaves but weakly expressed in pistils, shoot apices, and stamens (**Figure 3A**). When we analyzed the AS variant levels, we found a different pattern from that observed for the AS gene levels. A large proportion of AS variants were highly expressed in leaves but weakly expressed in pistils, shoot apices, and stamens (**Figure 3B**). A gene can produce multiple splice variants through AS, and the expression levels of

splice variants produced by the same gene can be different in different tissues (Sun and Xiao, 2015). Thus, the differences in the expression levels of AS genes and splice variants were unexpected.

Differentially Expressed SR Genes and hnRNP Genes

SR proteins and hnRNPs play important roles in regulating constitutive splicing and AS (Vitulo et al., 2014). To identify the SR genes and hnRNP genes that may affect AS, we analyzed the expression levels of such genes in seven tissues of *L. chinense*. Ultimately, we detected eight differently expressed SR genes and six hnRNP genes (**Figures 3C, D**). Intriguingly, all eight SR genes were expressed at low levels in leaves while two SR genes (*Lchi03086* and *Lchi06474*) were highly expressed in petals and sepals, five SR genes (*Lchi03662, Lchi04562, Lchi28029, Lchi06611,* and *Lchi22871*) were highly expressed in pistils, and six SR genes (*Lchi15805, Lchi03662, Lchi04562, Lchi28029, Lchi06611,* and *Lchi22871*) were highly expressed in shoot apices (**Figure 3C**). Among the six differentially expressed hnRNP genes, one (*Lchi16187*) was strongly expressed in leaves; the other five hnRNP genes (*Lchi12419, Lchi12420, Lchi29410, Lchi12536,* and *Lchi15250*) were weakly expressed in leaves, and four of them (*Lchi16187, Lchi29410, Lchi12536,* and *Lchi15250*) were also highly expressed in pistils and shoot apices (**Figure 3D**). Notably, most of the differentially expressed SR and hnRNP genes were expressed at low levels in leaves and at high levels in pistils and shoot apices; these findings were consistent with the expression patterns of most AS genes in leaves, pistils and petals (**Figures 3A, C, D**). We speculated that these differentially expressed SR and hnRNP genes may play important roles in AS regulation of *L. chinense*.

IR Affects AS Transcripts

As the most common AS event in plants, IR can make genes produce nonproductive transcripts by introducing PTCs into the coding sequence, and transcripts with PTCs can be degraded by the NMD pathway; thus, this process affects the proteome diversity (Marquez et al., 2012; Vitulo et al., 2014; Filichkin et al., 2018). To investigate how many PTC-containing (PTC$^+$) transcripts and non-PTC-containing (PTC$^-$) transcripts were produced due to IR, we analyzed the number of PTC$^+$ transcripts and PTC$^-$ transcripts in seven tissues. Firstly, we filtered the transcripts with IR that had FKPM values less than 1 in all tissues. Then, we identified a total of 1,656 PTC$^+$ transcripts, among which the majority were detected in leaves (1,132) and few were present in stamens (976) (**Figure 4A**). A total of 555 PTC$^+$ transcripts were present in all seven tissues, and 117, 64, 49, 36, 25, 23, and 15 tissue-specific PTC$^+$ transcripts were detected in leaves, stamens, shoot apex, pistils, petals, sepals, and bracts, respectively (**Supplementary Material 3, Figure S2A**). Moreover, we identified a total of 3,310 PTC$^-$ transcripts; the number of PTC$^-$ transcripts in seven tissues ranged from 2,190 (stamens) to 2,349 (leaves) (**Figure 4A**). Moreover, 1,383 PTC$^-$ transcripts existed in all seven tissues, and the 513 tissue-specific PTC$^-$ transcripts included 188 leaf-
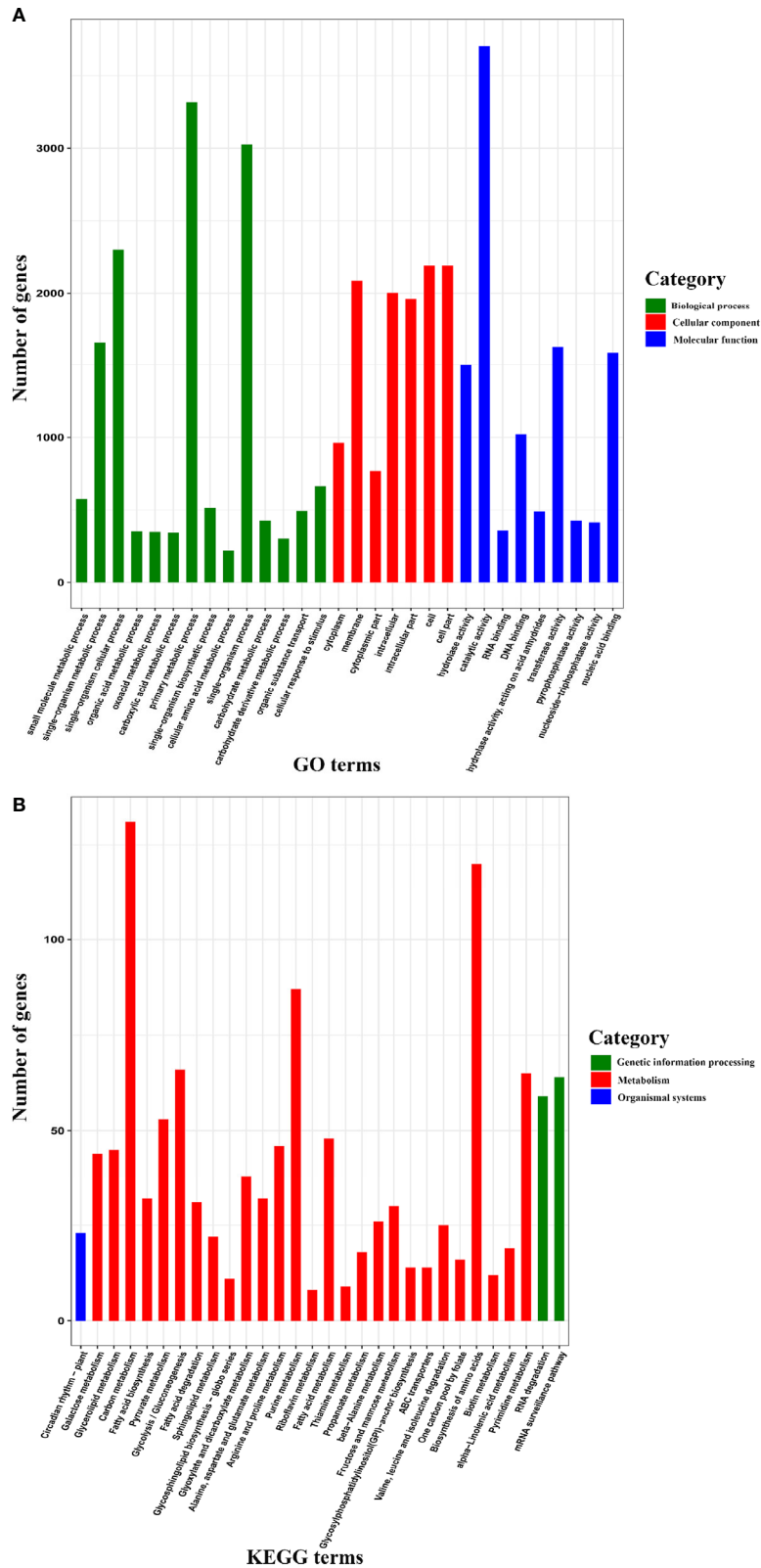
FIGURE 2 | Gene Ontology (GO) database and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of alternative splicing (AS) genes (only the top 30 terms are shown). **(A)** GO enrichment analysis of AS genes. **(B)** KEGG pathway enrichment analysis of AS genes.
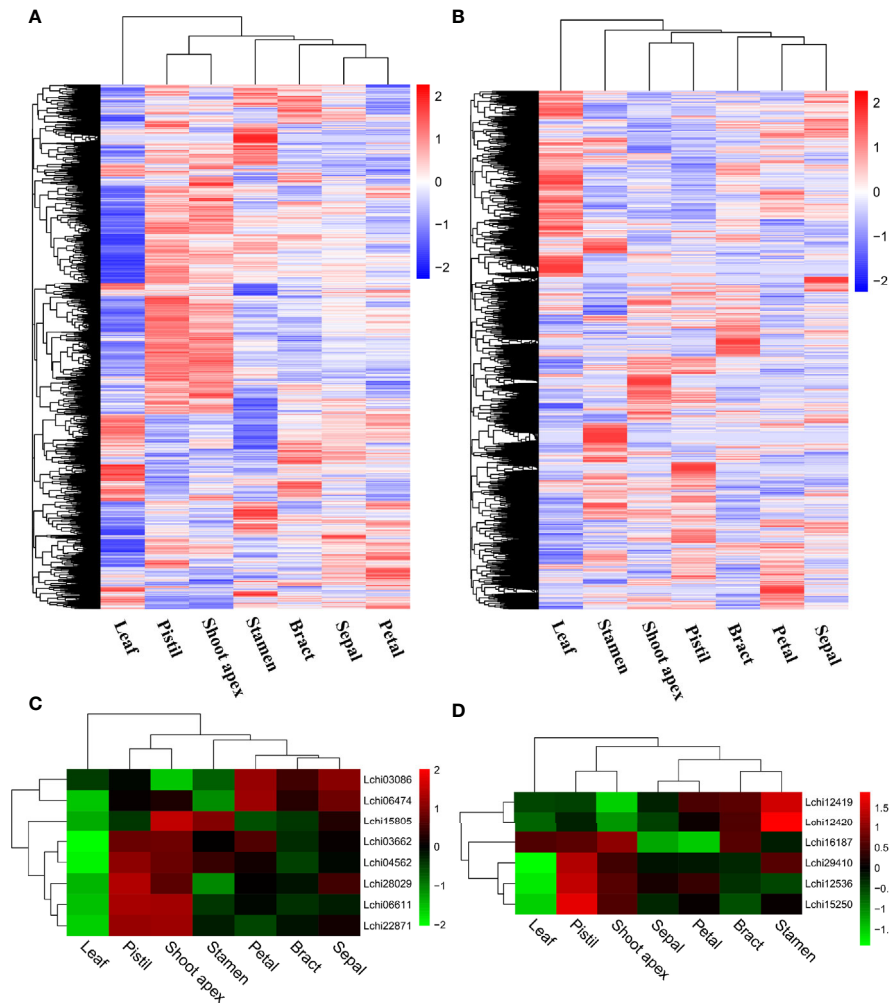
FIGURE 3 | Hierarchical clustering analysis and the relative expression profile of genes in different tissues of *L. chinense*. Hierarchical clustering analysis of alternative splicing (AS) genes **(A)** and alternatively spliced variants **(B)** in seven tissues. Differentially expressed serine/arginine-rich (SR) protein genes **(C)** and heterogeneous nuclear ribonucleoprotein (hnRNP) genes **(D)** in seven tissues.
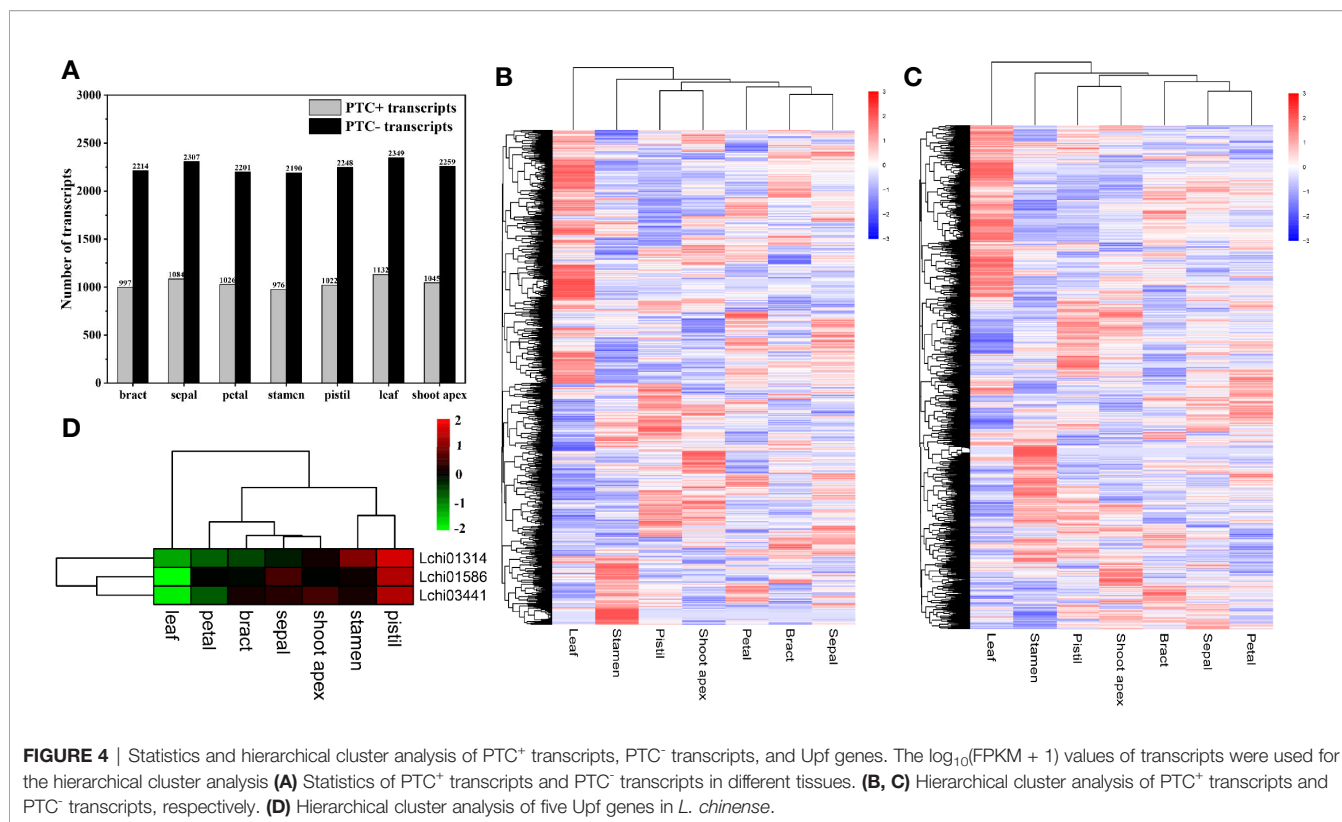
specific, 108 stamen-specific, 70 pistil-specific, 56 shoot-apex-specific, 33 bract-specific, 32 petal-specific, and 26 sepal-specific PTC$^-$ transcripts (**Supplementary Material 3, Figure S2B**). Overall, we found that the ratio between the PTC$^+$ transcript number and PTC$^-$ transcript number was approximately one to two (**Figure 4A**). Through the hierarchical cluster analysis, we found that a large number of transcripts, including both PTC$^+$ transcripts and PTC$^-$ transcripts, were highly expressed in leaves (**Figures 4B, C**).

Recent studies show that NMD regulates approximately 10% of mRNAs by downregulating them to meet the demands of organisms, and during this process three up-frameshift (Upf) proteins (Upf1, Upf2, and Upf3) play pivotal roles (Karousis et al., 2016; Gupta and Li, 2018). To determine the potential relationship between Upf proteins and PTC$^+$ transcripts at the expression level, we identified three Upf homologous genes (*Lchi01314* [Upf1], *Lchi03441* [Upf2], and *Lchi01586* [Upf3]) in *L. chinense* and confirmed their expression profile. We found

that these three Upf homologous genes tended to highly expressed in pistils, sepals, stamens, and shoot apices but lowly expressed in leaves, bracts, and petals (**Figure 4D**). Especially in pistils and leaves, the expression profile of the three Upf homologous genes in these two tissues was obviously opposite (**Figure 4D**). However, we could not find a stronger relationship between Upf genes and PTC$^+$ transcripts at the expression level. Notably, not all PTC$^+$ transcripts are targeted by NMD (Feng et al., 2015). We speculated that the low proportion of PTC$^+$ transcripts targeted by NMD may mask the relationship between Upf genes and PTC$^+$ transcripts at the expression level.

WGCNA Analysis of AS Genes and Hub Gene Identification

Although we found that most AS genes, SR genes, hnRNP genes, and Upf genes had opposite expression patterns between leaves and pistils, we could not identify the relationship between these

FIGURE 4 | Statistics and hierarchical cluster analysis of PTC⁺ transcripts, PTC⁻ transcripts, and Upf genes. The $\log_{10}$(FPKM + 1) values of transcripts were used for the hierarchical cluster analysis **(A)** Statistics of PTC⁺ transcripts and PTC⁻ transcripts in different tissues. **(B, C)** Hierarchical cluster analysis of PTC⁺ transcripts and PTC⁻ transcripts, respectively. **(D)** Hierarchical cluster analysis of five Upf genes in *L. chinense*.

patterns. To solve this problem, we performed WGCNA analysis of the AS genes, SR genes, hnRNP genes, and Upf genes. To improve the reliability of the analysis results, we selected the AS genes for which the rank of the variance of the FPKM value was in the top 60% for analysis. In total, 5,056 genes were classified into 6 coexpression modules, among which 1,141; 841; 534; 616; 1,147, and 777 AS genes were classified into the blue, brown, green, gray, turquoise, and yellow modules, respectively (**Figures 5A, B**). We also found that 70, 36, 13, 39, 45, and 70 AS transcription factors (TFs) were classified into the blue, brown, green, gray, turquoise, and yellow modules, respectively (**Figure 5C**). Then, we performed a hierarchical cluster analysis of these AS genes in the six coexpression modules. We found that AS genes in different modules showed obviously different expression patterns (**Figure 6**). Most AS genes from the blue module were strongly expressed in the pistils and shoot apices but weakly expressed in the leaves (**Figure 6A**). AS genes in the yellow module also showed similar expression patterns (**Figure 6F**). Moreover, AS genes in the brown module were highly expressed in the stamens, and AS genes from the gray module tended to be highly expressed in the bracts (**Figures 6B, D**). However, a large number of AS genes from the turquoise module showed high expression levels in the leaves but presented low expression levels in the stamens and pistils (**Figure 6E**). AS genes in the green module were clustered into two subclusters. One subcluster showed high expression patterns in the stamens, pistils, shoot apices, and bracts, while the other subcluster showed high expression patterns in the sepals, petals, and leaves (**Figure 6C**).

To determine what functions these AS genes play, we carried out GO and KEGG enrichment analysis. In the results of the GO enrichment analysis, we found that although the expression patterns of blue-module AS genes and yellow-module AS genes were very similar, there were obvious differences between them in the GO enrichment analysis results. For example, blue-module AS genes were significantly enriched in 12 cellular component terms, but yellow-module AS genes were not significantly enriched in cell component terms (**Supplementary Material 4, Figures S3A, F**). Moreover, blue-module AS genes were significantly enriched in transcription, post-transcription regulation, and translation related terms, such as translation, gene expression, ribosome biogenesis, RNA binding, DNA binding, and nucleic acid binding (**Supplementary Material 4, Figure S3A**). AS genes in the brown module were significantly enriched in carbohydrate metabolism-related terms, such as carbohydrate metabolic processes, hydrolase activity, and galactosidase activity (**Supplementary Material 4, Figure S3B**). AS genes from the green module were significantly enriched in multiple metabolic process (**Supplementary Material 4, Figure S3C**). However, AS genes in the gray module were only significantly enriched in six terms (**Supplementary Material 4, Figure S3D**). Moreover, AS genes in the turquoise module were significantly enriched in cytoplasm-related and oxidoreductase activity-related terms (**Supplementary Material 4, Figure S3E**). According to the KEGG pathway enrichment analysis results, only AS genes in the blue, brown, green, and turquoise modules were significantly
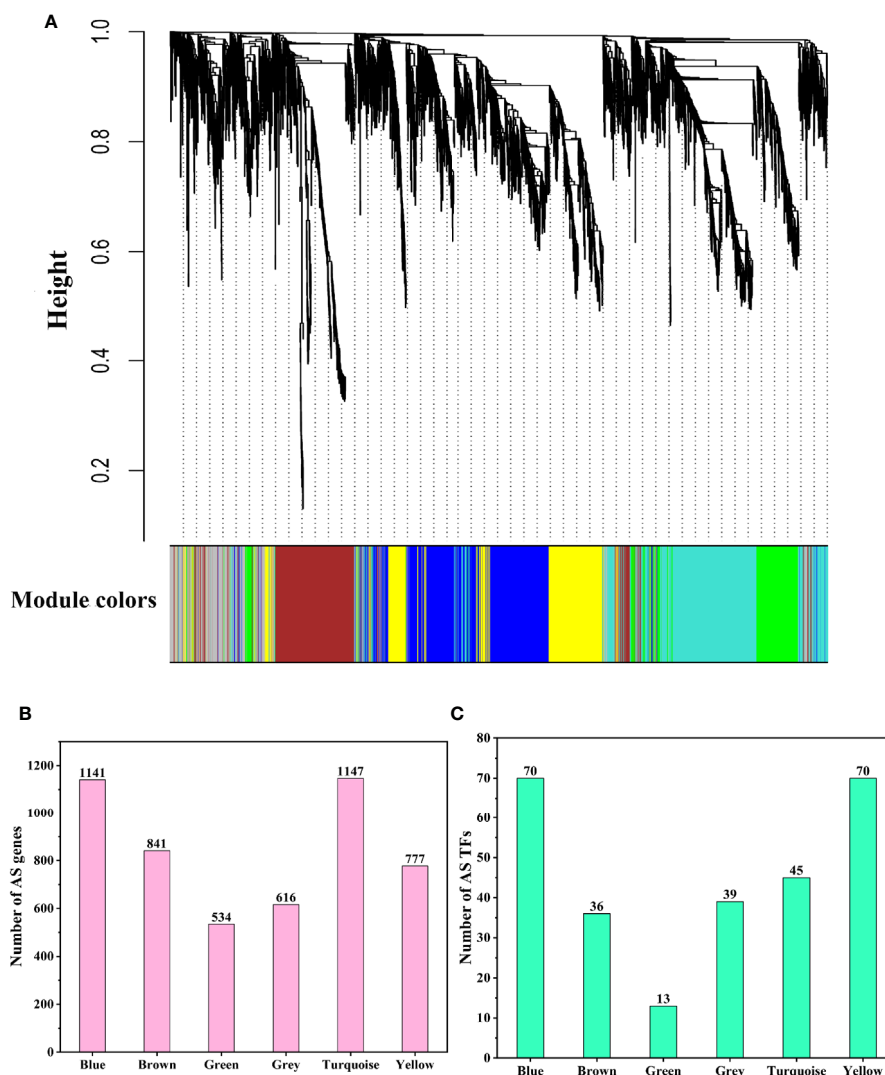
FIGURE 5 | Dendrogram of coexpressed alternative splicing (AS) genes and statistics of AS gene numbers and AS transcription factor (TF) numbers. **(A)** Dendrogram of coexpressed AS genes in *L. chinense*. The genes and coexpression modules are represented by vertical lines (upper panel) and colored rectangles (bottom panel), respectively. **(B)** Numbers of AS genes in six coexpression modules. **(C)** Numbers of AS TFs in six coexpression modules.

enriched (**Supplementary Material 5, Figure S4**). It seemed that blue-module AS genes participated in transcription, translation, and post-transcriptional regulation, because they were significantly enriched in RNA transport, RNA polymerase, spliceosome, ribosome, and mRMA surveillance pathways (**Supplementary Material 5, Figure S4A**). We also found that AS genes from the brown module and green module were significantly enriched in multiple metabolism pathways, such as sphingolipid metabolism, glycerolipid metabolism, pyruvate metabolism, and fatty acid metabolism (**Supplementary Material 5, Figures S4B, C**). Moreover, we determined that AS genes from the turquoise module participate in photosynthesis, because they were significantly enriched in the glyoxylate and dicarboxylate metabolism pathway and carbon metabolism pathway (**Supplementary Material 5, Figure S4E**).

To identify the hub genes that may play pivotal roles in different pathways, the degree algorithm in the CytoHubba package of Cytoscape software was applied. The networks of hub genes and their related genes in six coexpression modules were constructed and are shown in **Figure 7**. In the blue module, we identified 114 hub genes, among which 21 hub genes were involved in 7 significantly enriched KEGG pathways. *Lchi21288* (GTP binding protein) was related to the largest number of genes (1,121) (**Figure 7A**). Four AS TFs were identified as hub genes in the blue module, which were *Lchi01126* (FAR1 family), *Lchi29242* (C2H2 family), *Lchi05152* (C2H2-GATA family), and *Lchi21823* (C2H2 family) (**Figure 7A**). In the brown module, 84 hub genes were identified, including 5 AS TFs—*Lchi06210* (bZIP family), *Lchi27955* (MYB family), *Lchi26255* (MADS-M-type family), *Lchi11540* (bHLH family), and *Lchi17316* (NAC family)—and
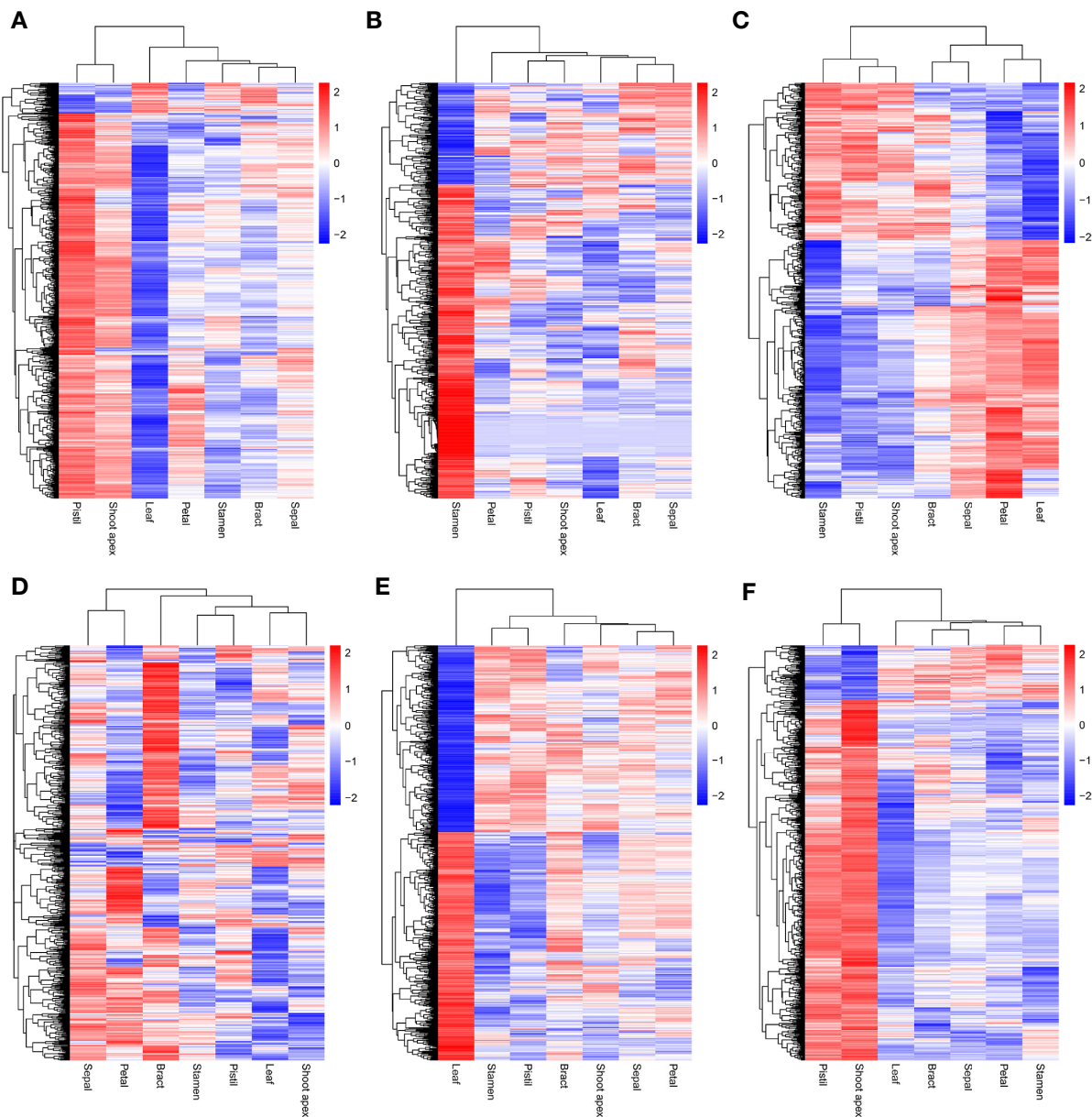
FIGURE 6 | Hierarchical cluster analysis of coexpressed alternative splicing (AS) genes in six modules. The $\log_{10}$(FPKM + 1) values of AS genes were used for the analysis. **(A–F)** Hierarchical cluster analysis of coexpressed AS genes in the blue, brown, green, gray, turquoise, and yellow modules, respectively.

Lchi12869 (Delta-1-pyrroline-5-carboxylate synthase), which were related to the largest number of genes (837). Notably, *Lchi10151* (Alpha-galactosidase 1) was involved in 4 significantly enriched KEGG pathways (**Figure 7B**). For the green module, 53 hub genes were identified, and *Lchi14809* (peptide chain release factor PrfB1) was associated with the largest number of genes (516); however, only 2 hub genes were participated in 1 significantly enriched KEGG pathway (**Figure 7C**). In the network of the gray module, 62 AS genes were considered as hub genes, including 2 AS TFs— *Lchi22536* (CAMTA family) and *Lchi00208* (CPP family)—and *Lchi30455* (LRR receptor-like serine/threonine-protein kinase),

which had a relationship with the maximum number of genes (613) (**Figure 7D**). For the turquoise module, we identified 115 hub genes, among which 18 hub genes participated in 3 significantly enriched KEGG pathways, among which *Lchi20763* (kynurenine–oxoglutarate transaminase 1-like protein), *Lchi29168* (epimerase family protein SDR39U1 homolog), and *Lchi13522* (photosynthetic NDH subunit of subcomplex B 5) were the top three hub genes related to 1,115, 1,084, and 1,073 genes, respectively (**Figure 7E**). In the yellow module network, 78 hub genes were identified, including 4 AS TFs—*Lchi00996* (MYB family), *Lchi07039* (C3H family), *Lchi08842* (MYB family), and
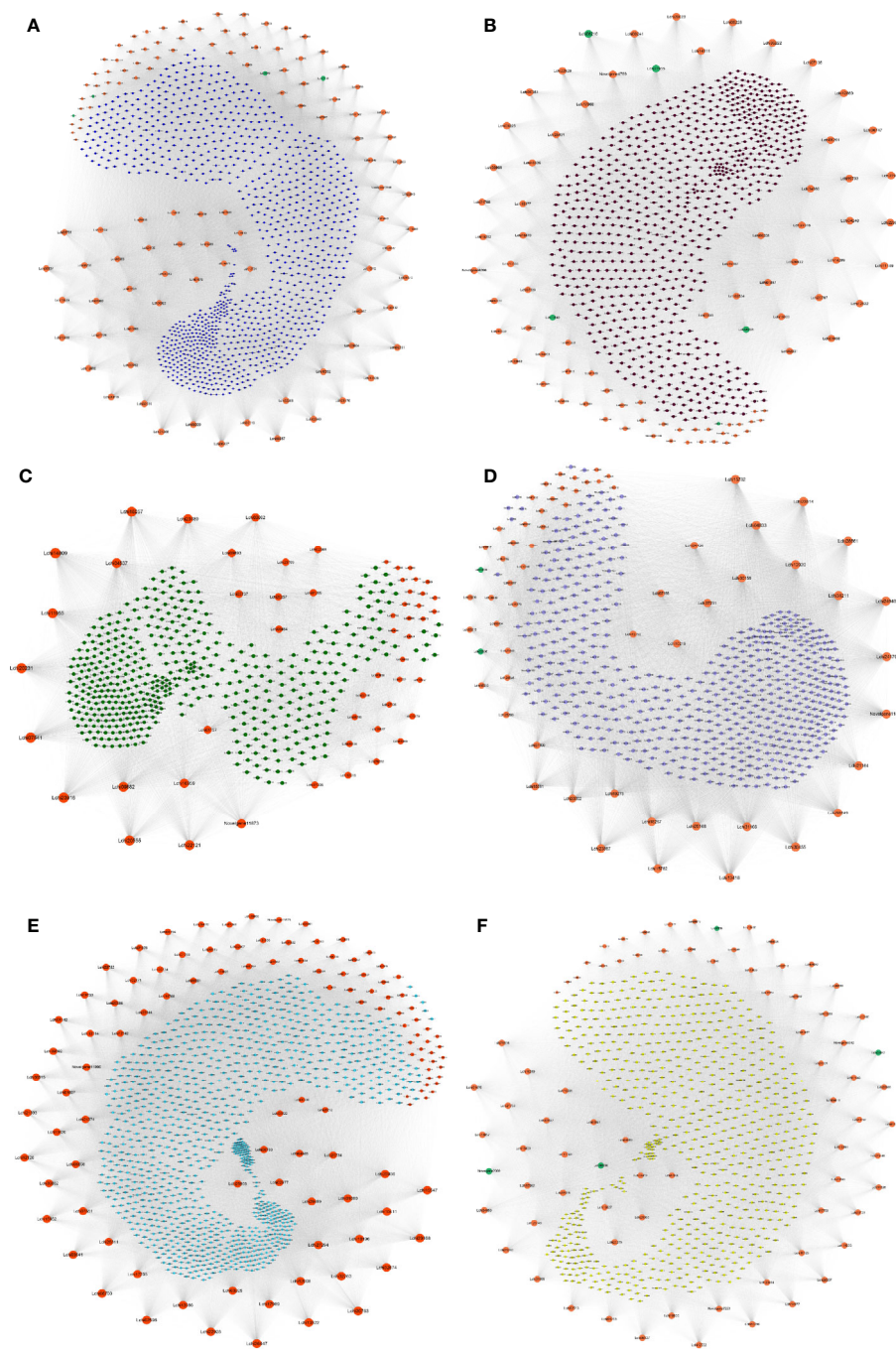
FIGURE 7 | Network of hub alternative splicing (AS) genes and their related genes in six coexpression modules. Hub genes are represented by orange-red notes, medium-aquamarine notes represent TFs that were also identified as hub genes, and other colored notes represent genes that were related to hub genes. The size of the note represents the degree of interaction between the hub gene and other genes, and larger notes indicate that the hub gene interacts with more genes. **(A–F)** Network of hub AS genes and their related genes in the blue, brown, green, gray, turquoise, and yellow modules, respectively.

Novelgene2009 (bHLH family)—and *Lchi34059* (protein IQ-DOMAIN), which was associated with the largest number of genes (752) (**Figure 7F**).

We could deduce from the above results that blue-module genes related to gene expression and post-transcriptional regulation are strongly expressed in pistils but weakly expressed in leaves, leading to differences in the activities of gene expression and post-transcriptional regulation between these two tissues and further resulting in genes related to AS, such as SR genes, hnRNP genes, and even the Upf gene, showing

opposite expression patterns in pistils and leaves. Moreover, we also found that AS affects multiple physiological and biochemical processes in *L. chinense*, such as photosynthesis, sphingolipid metabolism, fatty acid biosynthesis and metabolism. Furthermore, AS TFs were identified as hub genes in four coexpression modules, and it seemed that AS could influence *L. chinense* by influencing TFs.

Differences Between AS Genes and Non-AS Genes

To identify the differences between AS genes and non-AS genes, we compared the transcript lengths, exon numbers, total exon lengths, total intron lengths, average exon lengths and average intron lengths between AS genes and non-AS genes. We found

that AS genes tended to produce longer transcripts and have more exons compared to non-AS genes (**Figures 8A, B**). We also noticed that the total lengths of exons and introns of AS genes were longer than those of non-AS genes (**Figure 8C**). On average, the exons of AS genes were shorter than those of non-AS genes, while the introns of AS genes were longer than those of non-AS genes (**Figure 8D**). This indicates that the structural feature of genes may affect AS.

To investigate whether differences existed in splicing sites between AS genes and non-AS genes, we extracted two nucleotides from the ends of introns and analyzed the proportions of different splicing sites in the genes. We observed that for both AS genes and non-AS genes, canonical splicing sites (5′-GT.AG-3′) were predominant (70.46% for AS
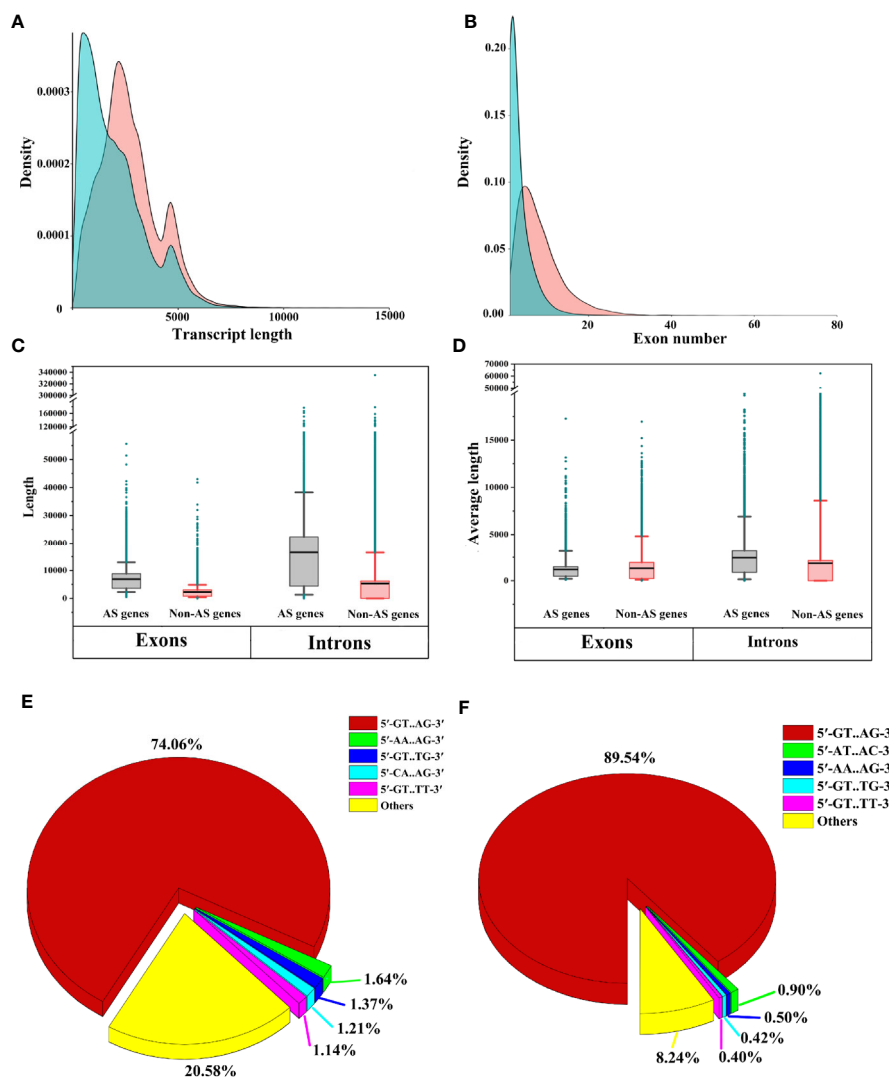

FIGURE 8 | Comparative analysis of alternative splicing (AS) genes and non-AS genes. **(A)** Comparison of transcript lengths between AS genes and non-AS genes. **(B)** Comparison of the numbers of exons of AS genes and non-AS genes. **(C)** Box plot of exon lengths and intron lengths of AS genes and non-AS genes. **(D)** Box plot of the average lengths of exons and introns of AS genes and non-AS genes. Boxes represent 25th and 75th percentiles. Whiskers represent 5th and 95th percentiles, and black lines in the boxes represent the median values. Dark cyan circles represent outliers. **(E)** Proportion of splicing sites in AS genes (only shown top 5 types). **(F)** Proportion of splicing sites in non-AS genes (only the top 5 types are shown).

genes and 89.54% for non-AS genes) (**Figures 8E, F**). However, noncanonical splicing sites, such as 5′-AA.AG-3′, 5′-GT.TG-3′, and 5′-GT.TT-3′, accounted for only small proportions of the splicing sites in AS genes and non-AS genes (**Figures 8E, F**). These results are consistent with findings obtained for tomato and *Ananas comosus* var. *bracteatus* (Clark et al., 2019; Ma et al., 2019).

Relationships Among AS Frequency, AS Events, and Gene Features

In this study, we identified 37,844 AS events from 8,503 AS genes (**Figure 1F**), indicating that multiple AS events occurred for a single gene. We next sought to determine whether the frequencies of AS events were related to gene features such as the numbers and lengths of exons/introns. Therefore, we compared the total lengths of exons/introns, the average lengths of exons/introns, and the numbers of exons/introns among AS genes with different AS event frequencies. We found that the total exon lengths and total intron lengths of AS genes increased with increasing AS frequencies (**Figures 9A, B**). However, when we compared the average lengths of exons/introns, we found that the relationship between the AS event frequency and the average exon/intron length was not strong (**Figures 9C, D**). In addition, the number of exons or introns increased with increasing AS event frequencies (**Figures 9E, F**). These findings suggest that genes with more complex structures tend to undergo more AS events.

To determine whether gene features were related to AS event types, we compared the average lengths of exons/introns and the numbers of exons among AS genes undergoing different types of AS events. We found that genes that underwent AL events had the longest average lengths of exons and introns (**Figures 10A, B**). Genes that underwent MX events or ES events had the shortest average exon lengths, and genes that underwent IR had the shortest average intron lengths (**Figures 10A, B**). Interestingly, genes that underwent AL events had the fewest exons, whereas genes that underwent MX events had the most exons (**Figure 10C**). These results indicate that shorter exons tend to be skipped, shorter introns tend to be retained, and genes with more exons tend to experience MX events. These phenomena are also observed in moso bamboo (Wang et al., 2018).

Identification of Alternatively Spliced lncRNAs

In this study, we identified 183 alternatively spliced lncRNAs (**Figure 10D**). Among these lncRNAs, 97 underwent A3 events, 88 underwent IR events, 77 underwent A5 events, 58 underwent ES events, 22 underwent AL events, 21 underwent AF events, and 8 underwent MX events (**Figure 10D**). A total of 972 AS events occurred in the 183 alternatively spliced lncRNAs (**Figure 10E**). IR events (258, 26.54%) constituted the largest proportion of AS events in lncRNAs, followed by AL (202, 20.78%), A3 (169, 17.39%), A5 (166, 17.08%), and ES (106, 10.91%) events, whereas MX events (16, 1.65%) constituted the smallest proportion of AS events in lncRNAs (**Figure 10E**). Through the hierarchical clustering analysis, we found that most alternatively spliced lncRNAs were preferentially highly expressed in shoot apices, stamens, and pistils (**Figure 10F**). Interestingly, the expression patterns of alternatively spliced lncRNAs appeared to be similar to those of the differently expressed SR proteins and hnRNP genes (**Figure 3**), indicating that SR proteins and hnRNP genes may substantially influence the production of alternatively spliced lncRNAs.

Validation of AS Events Using PCR

In this study, we randomly selected 8 AS genes and used PCR to validate the AS events. We identified 12 AS events that occurred in these genes. Among them, two events occurred in *Lchi20479*, one event occurred in *Lchi00028*, three events occurred in *Lchi27918*, one event occurred in *Lchi30260*, one event occurred in *Lchi00012*, one event occurred in *Lchi22958*, two events occurred in *Lchi24398*, and one event occurred in *Lchi19924* (**Figure 11**). The AS events that were validated by PCR were consistent with our sequencing results, indicating that our analysis results based on hybrid sequencing data were reliable.

DISCUSSION

AS is an important post-transcriptional regulation mechanism that enables one gene to produce multiple splice variants, thus increasing the complexity of the transcriptome and proteome (Lee and Rio, 2015; Hartmann et al., 2016). Moreover, AS plays pivotal roles in plant growth and development, the stress response, flowering, biological rhythms, and signal transduction (Staiger and Brown, 2013; Wang et al., 2014; Sun and Xiao, 2015; Dong et al., 2018). In *A. thaliana*, a splice variant of *auxin response factor 8* has been found to play a role in regulating endothecium lignification and stamen elongation (Ghelli et al., 2018). *Plant intracellular Ras-group leucine-rich repeat protein 6* is necessary for female and male gametogenesis in *A. thaliana*, and it is regulated by AS (Forsthoefel et al., 2018). Moreover, many studies on AS have been performed in different plant species, such as *A. thaliana*, tomato, moso bamboo, and peanut (Wang et al., 2014; Wang et al., 2017; Ruan et al., 2018; Zhao et al., 2018; Clark et al., 2019). However, specific studies on AS in *L. chinense* remain scarce. In this study, we combined SMRT sequencing with Illumina sequencing to obtain transcriptome data from 21 *L. chinense* samples. After data processing and analysis, we obtained large amounts of information about AS in *L. chinense*. This is the first systematic study of AS in this species.

To our knowledge, a large proportion of genes undergo AS events in plants. In *A. thaliana*, more than 61% of multi-exon genes undergo AS events (Marquez et al., 2012). In maize, peanut, moso bamboo, and tomato, the proportions of AS genes are approximately 55%, 37%, 49%, and 65%, respectively (Min, 2017; Ruan et al., 2018; Zhao et al., 2018; Clark et al., 2019). Due to the limited sample size and sequencing depth in this study, we found that only 23.66% (8,293/35,055) of multi-exon
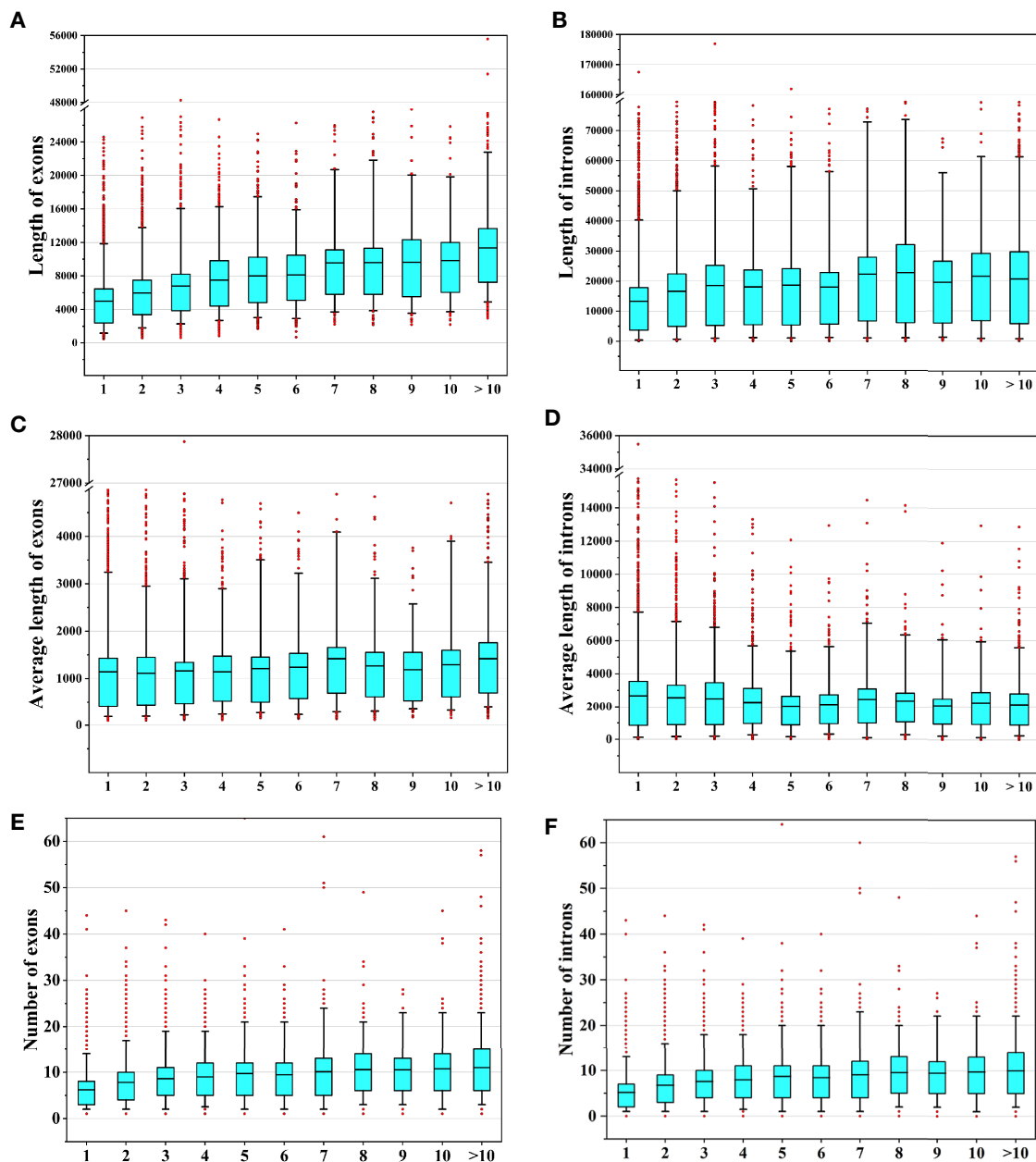
FIGURE 9 | Analysis of alternative splicing (AS) events and gene features. **(A, B)** Comparison of the total lengths of exons/introns of AS genes that underwent different numbers of AS events. **(C, D)** Comparison of the average lengths of exons/introns of AS genes that underwent different numbers of AS events. **(E, F)** Comparison of the numbers of exons/introns of AS genes that underwent different numbers of AS events.

genes in *L. chinense* were alternatively spliced (**Figure 1**). Thus, the proportion of alternatively spliced multi-exon genes in *L. chinense* is apparently lower than the proportions in the above species. Notably, however, the number of samples and the quality and depth of sequencing have been found to be positively correlated with the proportion of alternatively spliced multi-exon genes (Ruan et al., 2018). For example, with improvements in the quality and depth of sequencing, the proportion of AS genes has been found to increase from 22%

to greater than 61% in *A. thaliana* (Reddy, 2007; Marquez et al., 2012). In moso bamboo, the AS rate has increased from ~36% to greater than 49%, mainly because of increases in sample size (Wang et al., 2018; Zhao et al., 2018). We believe that the proportion of known AS genes in *L. chinense* will increase with increasing sample size and sequencing depth.

Unlike in animals, the most common AS event in plants is IR rather than ES (Chaudhary et al., 2019b). In this study, IR (10,743, 28.39%) accounted for the highest proportion of AS
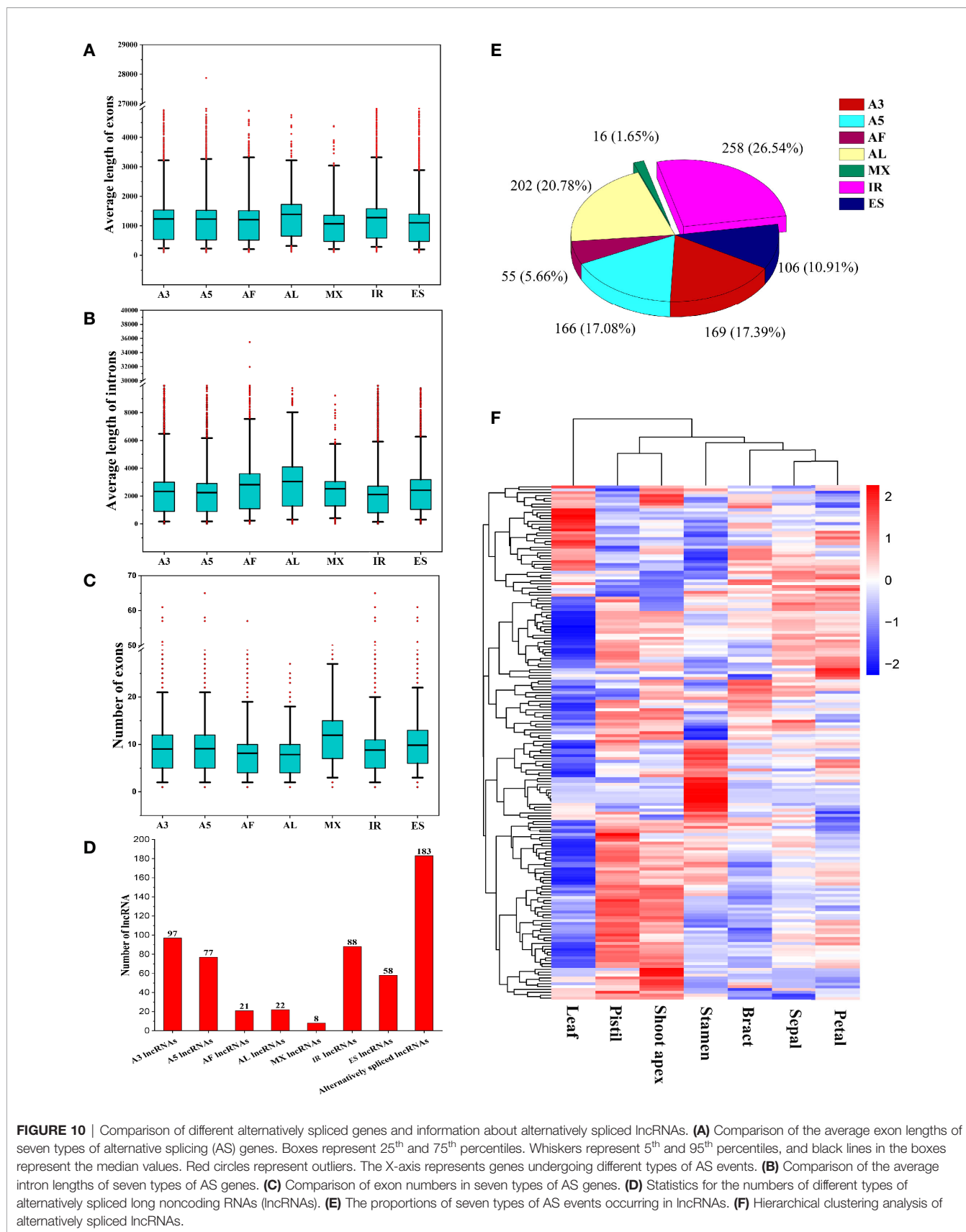
FIGURE 10 | Comparison of different alternatively spliced genes and information about alternatively spliced lncRNAs. **(A)** Comparison of the average exon lengths of seven types of alternative splicing (AS) genes. Boxes represent 25th and 75th percentiles. Whiskers represent 5th and 95th percentiles, and black lines in the boxes represent the median values. Red circles represent outliers. The X-axis represents genes undergoing different types of AS events. **(B)** Comparison of the average intron lengths of seven types of AS genes. **(C)** Comparison of exon numbers in seven types of AS genes. **(D)** Statistics for the numbers of different types of alternatively spliced long noncoding RNAs (lncRNAs). **(E)** The proportions of seven types of AS events occurring in lncRNAs. **(F)** Hierarchical clustering analysis of alternatively spliced lncRNAs.
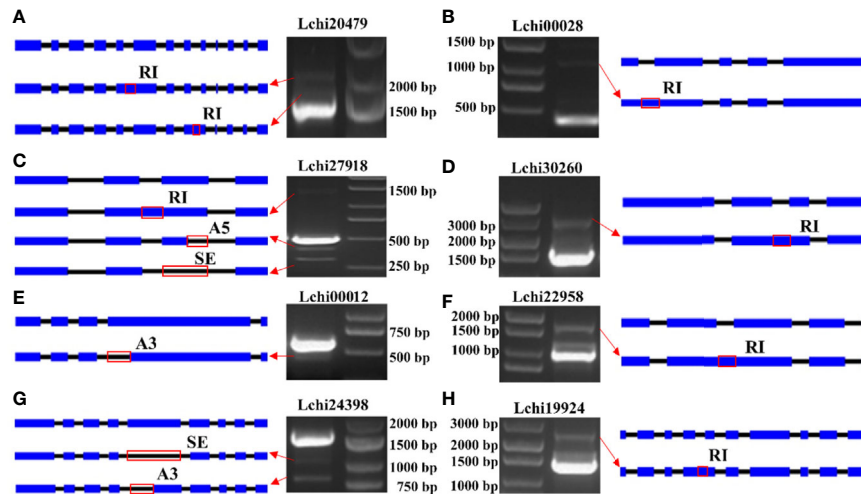
FIGURE 11 | Validation of alternative splicing (AS) events in eight AS genes. **(A–H)** Gene structure and polymerase chain reaction (PCR) results of *Lchi20479*, *Lchi00028*, *Lchi27918*, *Lchi30260*, *Lchi00012*, *Lchi22958*, *Lchi24398*, and *Lchi19924*. Blue rectangles represent exons, black lines represent introns, and red rectangles represent the positions where the event occurred.

events (**Figure 1F**). This finding is consistent with those obtained in studies on tomato, *A. thaliana*, and maize (Marquez et al., 2012; Min, 2017; Clark et al., 2019). Moreover, ES (6,312, 16.68%) was the third largest AS event (**Figure 1F**). Both ES and IR contribute to increasing the complexity of plant transcriptomes and proteomes. ES manifests its influence by affecting protein-protein interaction sites and post-translational modification sites (Buljan et al., 2012; Ellis et al., 2012). IR occurs in open reading frames (ORFs), affecting transcriptome and proteome complexity in two ways—one is by introducing PTCs into the ORF to trigger NMD or produce truncated proteins, and the other is by producing an elongated ORF that allows the production of pairs of protein isoforms (Jacob and Smith, 2017). Moreover, we identified 1,656 PTC+ transcripts and 3,310 PTC- transcripts (**Figure 4A**). Actually, these 1,656 PTC+ transcripts may only partially trigger NMD, as a previous analysis showed that approximately one third of the PTC+ transcripts are likely to trigger NMD (Lewis et al., 2003). In addition, NMD effectors also play important roles in the degradation of PTC+ transcripts through the NMD pathway, including Upf proteins. Although we found that the expression patterns of the three Upf genes in pistils and leaves were opposite, we did not find any obvious relationship between Upf genes and PTC+ transcripts. However, we still believe that AS causes genes to produce multiple splice variants, substantially increasing the complexity and diversity of the *L. chinense* transcriptome and proteome.

Tissue-specific AS is controlled by SFs (Chen and Manley, 2009). SR proteins and hnRNPs are the main SFs that play roles in regulating AS (Vitulo et al., 2014). These two SFs play roles in different manners; SR proteins regulate AS by participating in spliceosome formation while hnRNPs usually act as repressors or activators (Mauger et al., 2008; Mai et al., 2016). Many SR

proteins and hnRNPs have been reported to be involved in AS regulation, such as SRp38, SRp40, RBM39, hnRNP A1, hnRNP F, and hnRNP H, (Feng et al., 2009; Yan et al., 2010; Talukdar et al., 2011; Mai et al., 2016; Nazim et al., 2017). Differential expression of SR proteins and hnRNPs affects AS in different tissues (Chen and Manley, 2009; Vitulo et al., 2014). In grapes, differentially expressed SR proteins are related to tissue-specific AS (Vitulo et al., 2014). We found that 13 of the 14 differentially expressed SR and hnRNP genes showed opposite expression patterns in leaves and pistils, and their expression patterns were consistent with those of most AS genes (**Figure 3**). Furthermore, three Upf genes that play critical roles in the NMD pathway also showed an opposite expression pattern in these two tissues (**Figure 4D**). We speculate that differentially expressed SR and hnRNP genes and Upf genes may be related to the differences between pistils and leaves at the transcriptome level, further impacting the complexity of the transcriptome.

To further reveal the potential roles that AS may play in *L. chinense*, we performed WGCNA of the AS genes. As a powerful tool with which to reveal the relationships between different gene sets or between genes and phenotypes, WGCNA has been extensively applied in different fields, such as medicine and plant science (Wang J.-L. et al., 2019; Zhu et al., 2019; Esmaeili et al., 2020; Jia et al., 2020). We found that pistils and leaves may have huge differences in transcription and post-transcriptional regulation, because most blue-module genes that were related to transcription and post-transcriptional regulation were strongly expressed in pistils but weakly expressed in leaves (**Figures 6A**, **S3A**, **and S4A**). AS, as an important post-transcriptional regulation mechanism, also makes contributions to these differences, so AS-related SR genes, hnRNP genes, and Upf genes also presented different expression patterns in pistils and leaves. Moreover, AS genes from different coexpression modules

participated in multiple physiological and biochemical processes, such as fatty acid biosynthesis and metabolism, galactose metabolism, and photosynthesis, and these processes were necessary for plant life. Furthermore, 15 AS TFs were identified as hub genes in four modules. Therefore, the effects of AS on *L. chinense* were varied, and these findings have broadened our view on AS of *L. chinense*.

To identify the differences between AS genes and non-AS genes, we compared these gene types and found that AS genes had more exons and introns and produced longer transcripts than did non-AS genes (**Figures 10A, B**). The number of exons/introns was positively related to the number of AS events (**Figures 9E, F**). These results indicate that genes with more complex structures tend to experience more AS events, which is consistent with the findings obtained in study on moso bamboo (Wang et al., 2018). However, there were no significant differences in splicing sites between AS genes and non-AS genes (**Figures 8E, F**). Similarly, in tomato, moso bamboo, and *A. comosus* var. *bracteatus*, researchers have reported that AS genes and non-AS genes have no differences in splicing site selection (Zhao et al., 2018; Clark et al., 2019; Ma et al., 2019). Moreover, we found that short exons tended to be skipped, while short introns tended to be retained (**Figures 10A, B**). The same phenomenon has been observed in tomatoes (Clark et al., 2019). We believe that the structural features of genes may affect AS. These findings greatly increase our understanding of AS in *L. chinense*.

AS can not only increase the complexity of the transcriptome and proteome but also influence lncRNAs (Kiegle et al., 2018; Yang et al., 2019). LncRNAs are critical regulators of plant growth and development, flowering, the stress response, and other biological processes (Liu et al., 2018; Ayachit et al., 2019). In grapevine, researchers have identified 813 differentially expressed lncRNAs related to the cold stress response (Wang P. et al., 2019). Moreover, researchers have found that more than 72% of lncRNAs in tomatoes undergo AS events (Yang et al., 2019). In our study, we identified 183 lncRNAs associated with AS (**Figure 10D**). We believe that the different sequencing and analysis methods used in our study may have restricted our detection of alternatively spliced lncRNAs in *L. chinense*. However, our findings still expand the current understanding of AS in *L. chinense*.

CONCLUSION

In this study, we used different tools and algorithms to effectively obtain information about AS in *L. chinense* from hybrid sequencing data. We identified a total of 48,408 genes, 8,503 of which were found to undergo AS. IR events (10,743, 28.39%) were the most common event, resulting in the production of 1,656 PTC$^+$ and 3,310 PTC$^-$ transcripts. Furthermore, we detected 183 lncRNAs that were alternatively spliced. These findings indicate that AS plays a critical role in increasing the complexity of the transcriptome of *L. chinense*. Moreover, WGCNA revealed that pistils and leaves had great differences in the activities of transcription and post-transcriptional regulation, and AS had an impact on many physiological and biochemical processes in *L. chinense*. We also found that structural features of genes may affect

AS. AS genes and non-AS genes had differences in transcript length, exon number, and exon/intron length, and AS genes with more exons or introns experienced more AS events. Furthermore, shorter introns tended to be retained, while shorter exons tended to be skipped. To validate our results, we randomly selected 8 AS genes for PCR, and the PCR results showed that our analysis was reliable. Our study provided new information about AS in *L. chinense*, and it is the first to comprehensively report on AS in this species. Our data thus provide a reference for further study of AS in *L. chinense*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA559687.

AUTHOR CONTRIBUTIONS

Experimental design: ZT and HL. Plant material collection and execution of the experiments: ZT, YS, SW, and YZ. Data analysis: ZT, YS, and SW. Manuscript writing: ZT and YZ. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.578100/full#supplementary-material

REFERENCES

Abdel-Ghany, S. E., Hamilton, M., Jacobi, J. L., Ngam, P., Devitt, N., Schilkey, F., et al. (2016). A survey of the sorghum transcriptome using single-molecule long reads. *Nat. Commun.* 7, 1–11. doi: 10.1038/ncomms11706

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Ayachit, G., Shaikh, I., Sharma, P., Jani, B., Shukla, L., Sharma, P., et al. (2019). De novo transcriptome of *Gymnema sylvestre* identified putative lncRNA and genes regulating terpenoid biosynthesis pathway. *Sci. Rep.* 9 (1), 1–13. doi: 10.1038/s41598-019-51355-x

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12 (1), 59–60. doi: 10.1038/nmeth.3176

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G. P., Fuxreiter, M., Bateman, A., et al. (2012). Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* 46 (6), 871–883. doi: 10.1016/j.molcel.2012.05.039

Calixto, C. P. G., Tzioutziou, N. A., James, A. B., Hornyik, C., Guo, W., Zhang, R., et al. (2019). Cold-Dependent Expression and Alternative Splicing of Arabidopsis Long Non-coding RNAs. *Front. Plant Sci.* 10:235:235. doi: 10.3389/fpls.2019.00235

Capitao, C., Shukla, N., Wandrolova, A., Mittelsten Scheid, O., and Riha, K. (2018). Functional Characterization of SMG7 Paralogs in Arabidopsis thaliana. *Front. Plant Sci.* 9:1602:1602. doi: 10.3389/fpls.2018.01602

Chaudhary, S., Jabre, I., Reddy, A. S. N., Staiger, D., and Syed, N. H. (2019a). Perspective on Alternative Splicing and Proteome Complexity in Plants. *Trends Plant Sci.* 24 (6), 496–506. doi: 10.1016/j.tplants.2019.02.006

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A. S. N., Byrne, L. J., Wilson, C. M., et al. (2019b). Alternative Splicing and Protein Diversity: Plants Versus Animals. *Front. Plant Sci.* 10:708:708. doi: 10.3389/fpls.2019.00708

Chen, M., and Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10 (11), 741–754. doi: 10.1038/nrm2777

Chen, C., Zheng, Y., Zhong, Y., Wu, Y., Li, Z., Xu, L. A., et al. (2018). Transcriptome analysis and identification of genes related to terpenoid biosynthesis in *Cinnamomum camphora*. *BMC Genomics* 19 (1), 550. doi: 10.1186/s12864-018-4941-1

Chen, J., Hao, Z., Guang, X., Zhao, C., Wang, P., Xue, L., et al. (2018). Liriodendron genome sheds light on angiosperm phylogeny and species-pair differentiation. *Nat. Plants* 5 (3), 18–25. doi: 10.1038/s41477-018-0323-6

Clark, S., Yu, F., Gu, L., and Min, X. J. (2019). Expanding Alternative Splicing Identification by Integrating Multiple Sources of Transcription Data in Tomato. *Front. Plant Sci.* 10, 689. doi: 10.3389/fpls.2019.00689

Dong, C., He, F., Berkowitz, O., Liu, J., Cao, P., Tang, M., et al. (2018). Alternative Splicing Plays a Critical Role in Maintaining Mineral Nutrient Homeostasis in Rice (*Oryza sativa*). *Plant Cell* 30 (10), 2267–2285. doi: 10.1105/tpc.18.00051

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986

Ellis, J. D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J. A., et al. (2012). Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol. Cell* 46 (6), 884–892. doi: 10.1016/j.molcel.2012.05.037

Esmaeili, S., Mehrgou, A., Kakavandi, N., and Rahmati, Y. (2020). Exploring Kawasaki disease-specific hub genes revealing a striking similarity of expression profile to bacterial infections using weighted gene co-expression network analysis (WGCNA) and co-expression modules identification tool (CEMiTool): An integrated bioinformatics and experimental study. *Immunobiology* 225 (4), 151980. doi: 10.1016/j.imbio.2020.151980

Feng, Y., Valley, M. T., Lazar, J., Yang, A. L., Bronson, R. T., Firestein, S., et al. (2009). SRp38 regulates alternative splicing and is required for Ca(2+) handling in the embryonic heart. *Dev. Cell* 16 (4), 528–538. doi: 10.1016/j.devcel.2009.02.009

Feng, D., Su, R. C., Zou, L., Triggs-Raine, B., Huang, S., and Xie, J. (2015). Increase of a group of PTC(+) transcripts by curcumin through inhibition of the NMD

pathway. *Biochim. Biophys. Acta* 1849 (8), 1104–1115. doi: 10.1016/j.bbagrm.2015.04.002

Filichkin, S., Priest, H. D., Megraw, M., and Mockler, T. C. (2015). Alternative splicing in plants: directing traffic at the crossroads of adaptation and environmental stress. *Curr. Opin. Plant Biol.* 24, 125–135. doi: 10.1016/j.pbi.2015.02.008

Filichkin, S. A., Hamilton, M., Dharmawardhana, P. D., Singh, S. K., Sullivan, C., Ben-Hur, A., et al. (2018). Abiotic Stresses Modulate Landscape of Poplar Transcriptome via Alternative Splicing, Differential Intron Retention, and Isoform Ratio Switching. *Front. Plant Sci.* 9:5:5. doi: 10.3389/fpls.2018.00005

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39 (Web Server issue), W29–W37. doi: 10.1093/nar/gkr367

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44 (D1), D279–D285. doi: 10.1093/nar/gkv1344

Forsthoefel, N. R., Klag, K. A., McNichol, S. R., Arnold, C. E., Vernon, C. R., Wood, W. W., et al. (2018). Arabidopsis PIRL6 Is Essential for Male and Female Gametogenesis and Is Regulated by Alternative Splicing. *Plant Physiol.* 178 (3), 1154–1169. doi: 10.1104/pp.18.00329

Gallegos, J. (2018). Alternative Splicing Plays a Major Role in Plant Response to Cold Temperatures. *Plant Cell* 30 (7), 1378–1379. doi: 10.1105/tpc.18.00430

Ghelli, R., Brunetti, P., Napoli, N., De Paolis, A., Cecchetti, V., Tsuge, T., et al. (2018). A Newly Identified Flower-Specific Splice Variant of AUXIN RESPONSE FACTOR8 Regulates Stamen Elongation and Endothecium Lignification in Arabidopsis. *Plant Cell* 30 (3), 620–637. doi: 10.1105/tpc.17.00840

Gupta, P., and Li, Y. R. (2018). Upf proteins: highly conserved factors involved in nonsense mRNA mediated decay. *Mol. Biol. Rep.* 45 (1), 39–55. doi: 10.1007/s11033-017-4139-7

Hao, R., He, S., Tang, S., and Wu, S. (1995). Geographical distribution of *Liriodendron chinense* in China and its significance. *J. Plant Resour. Environ.* 4 (4), 1–6.

Hartmann, L., Drewe-Boss, P., Wiessner, T., Wagner, G., Geue, S., Lee, H. C., et al. (2016). Alternative Splicing Substantially Diversifies the Transcriptome during Early Photomorphogenesis and Correlates with the Energy Availability in Arabidopsis. *Plant Cell* 28 (11), 2715–2734. doi: 10.1105/tpc.16.00508

Jacob, A. G., and Smith, C. W. J. (2017). Intron retention as a component of regulated gene expression programs. *Hum. Genet.* 136 (9), 1043–1057. doi: 10.1007/s00439-017-1791-x

Jia, R., Zhao, H., and Jia, M. (2020). Identification of co-expression modules and potential biomarkers of breast cancer by WGCNA. *Gene* 750, 144757. doi: 10.1016/j.gene.2020.144757

Kanehisa, M., and Goto, S. (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi: 10.1093/nar/28.1.27

Karlik, E., and Gozukirmizi, N. (2018). Expression analysis of lncRNA AK370814 involved in the barley vitamin B6 salvage pathway under salinity. *Mol. Biol. Rep.* 45 (6), 1597–1609. doi: 10.1007/s11033-018-4289-2

Karousis, E. D., Nasif, S., and Muhlemann, O. (2016). Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley Interdiscip. Rev. RNA* 7 (5), 661–682. doi: 10.1002/wrna.1357

Kiegle, E. A., Garden, A., Lacchini, E., and Kater, M. M. (2018). A Genomic View of Alternative Splicing of Long Non-coding RNAs during Rice Seed Development Reveals Extensive Splicing and lncRNA Gene Families. *Front. Plant Sci.* 9, 115. doi: 10.3389/fpls.2018.00115

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., et al. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35 (Web Server issue), W345–W349. doi: 10.1093/nar/gkm391

Laloum, T., Martin, G., and Duque, P. (2018). Alternative Splicing Control of Abiotic Stress Responses. *Trends Plant Sci.* 23 (2), 140–150. doi: 10.1016/j.tplants.2017.09.019

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Lee, Y., and Rio, D. C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* 84, 291–323. doi: 10.1146/annurev-biochem-060614-034316

Lewis, B. P., Green, R. E., and Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsensemediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* 100, 189–192. doi: 10.1073/pnas.0136770100

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22 (13), 1658–1659. doi: 10.1093/bioinformatics/btl158

Li, A., Zhang, J., and Zhou, Z. (2014). PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinf.* 15, 311–321. doi: 10.1186/1471-2105-15-311

Liu, S., Li, A., Chen, C., Cai, G., Zhang, L., Guo, C., et al. (2017). De Novo Transcriptome Sequencing in *Passiflora edulis* Sims to Identify Genes and Signaling Pathways Involved in Cold Tolerance. *Forests* 8 (11), 435. doi: 10.3390/f8110435

Liu, S., Sun, Z., and Xu, M. (2018). Identification and characterization of long non-coding RNAs involved in the formation and development of poplar adventitious roots. *Ind. Crops Prod.* 118, 334–346. doi: 10.1016/j.indcrop.2018.03.071

Lloyd, J. P., and Davies, B. (2013). SMG1 is an ancient nonsense-mediated mRNA decay effector. *Plant J.* 76 (5), 800–810. doi: 10.1111/tpj.12329

Ma, J., Xiang, Y., Xiong, Y., Lin, Z., Xue, Y., Mao, M., et al. (2019). SMRT sequencing analysis reveals the full-length transcripts and alternative splicing patterns in *Ananas comosus* var. *bracteatus*. *PeerJ* 7, e7062. doi: 10.7717/peerj.7062

Mai, S., Qu, X., Li, P., Ma, Q., Cao, C., and Liu, X. (2016). Global regulation of alternative RNA splicing by the SR-rich protein RBM39. *Biochim. Biophys. Acta* 1859 (8), 1014–1024. doi: 10.1016/j.bbagrm.2016.06.007

Marquez, Y., Brown, J. W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* 22 (6), 1184–1195. doi: 10.1101/gr.134106.111

Mauger, D. M., Lin, C., and Garcia-Blanco, M. A. (2008). hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc. *Mol. Cell Biol.* 28 (17), 5403–5419. doi: 10.1128/MCB.00739-08

Min, X. J. (2017). Comprehensive Cataloging and Analysis of Alternative Splicing in Maize. *Comput. Mol. Biol.* 7 (1), 1–11. doi: 10.5376/cmb.2017.07.0001

Nazim, M., Masuda, A., Rahman, M. A., Nasrin, F., Takeda, J. I., Ohe, K., et al. (2017). Competitive regulation of alternative splicing and alternative polyadenylation by hnRNP H and CstF64 determines acetylcholinesterase isoforms. *Nucleic Acids Res.* 45 (3), 1455–1468. doi: 10.1093/nar/gkw823

Park, Y. J., Lee, J. H., Kim, J. Y., and Park, C. M. (2019). Alternative RNA Splicing Expands the Developmental Plasticity of Flowering Transition. *Front. Plant Sci.* 10:606:606. doi: 10.3389/fpls.2019.00606

Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 35 (Database issue), D61–D65. doi: 10.1093/nar/gkl842

Reddy, A. S. (2007). Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.* 58, 267–294. doi: 10.1146/annurev.arplant.58.032806.103754

Riehs, N., Akimcheva, S., Puizina, J., Bulankova, P., Idol, R. A., Siroky, J., et al. (2008). Arabidopsis SMG7 protein is required for exit from meiosis. *J. Cell Sci.* 121 (Pt 13), 2208–2216. doi: 10.1242/jcs.027862

Ruan, J., Guo, F., Wang, Y., Li, X., Wan, S., Shan, L., et al. (2018). Transcriptome analysis of alternative splicing in peanut (*Arachis hypogaea* L.). *BMC Plant Biol.* 18 (1), 139. doi: 10.1186/s12870-018-1339-9

Salmela, L., and Rivals, E. (2014). LoRDEC: accurate and efficient long read error correction. *Bioinformatics* 30 (24), 3506–3514. doi: 10.1093/bioinformatics/btu538

Shen, Y., Zhou, Z., Wang, Z., Li, W., Fang, C., Wu, M., et al. (2014). Global dissection of alternative splicing in paleopolyploid soybean. *Plant Cell* 26 (3), 996–1008. doi: 10.1105/tpc.114.122739

Sreya, G., and Chon-Kit Kenneth, C. (2016). Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods Mol. Biol.* 1374, 339–361. doi: 10.1007/978-1-4939-3167-5_18

Staiger, D., and Brown, J. W. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* 25 (10), 3640–3656. doi: 10.1105/tpc.113.113803

Sun, Y., and Xiao, H. (2015). Identification of alternative splicing events by RNA sequencing in early growth tomato fruits. *BMC Genomics* 16, 948. doi: 10.1186/s12864-015-2128-6

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41 (17), e166–e173. doi: 10.1093/nar/gkt646

Szakonyi, D., and Duque, P. (2018). Alternative Splicing as a Regulator of Early Plant Development. *Front. Plant Sci.* 9, 1174. doi: 10.3389/fpls.2018.01174

Talukdar, I., Sen, S., Urbano, R., Thompson, J., Yates, J. R.,. 3., and Webster, N. J. (2011). hnRNP A1 and hnRNP F modulate the alternative splicing of exon 11 of the insulin receptor gene. *PloS One* 6 (11), e27869. doi: 10.1371/journal.pone.0027869

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinf.* 4, 41–54. doi: 10.1186/1471-2105-4-41

The UniProt, C. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45 (D1), D158–D169. doi: 10.1093/nar/gkw1099

Tognacca, R. S., Servi, L., Hernando, C. E., Saura-Sanchez, M., Yanovsky, M. J., Petrillo, E., et al. (2019). Alternative Splicing Regulation During Light-Induced Germination of *Arabidopsis thaliana* Seeds. *Front. Plant Sci.* 10, 1076. doi: 10.3389/fpls.2019.01076

Tu, Z., Hao, Z., Zhong, W., and Li, H. (2019). Identification of Suitable Reference Genes for RT-qPCR Assays in Liriodendron chinense (Hemsl.) Sarg. *Forests* 10 (5), 441–456. doi: 10.3390/f10050441

Vitulo, N., Forcato, C., Carpinelli, E. C., Telatin, A., Campagna, D., D'Angelo, M., et al. (2014). A deep survey of alternative splicing in grape reveals changes in the splicing machinery related to tissue, stress condition and genotype. *BMC Plant Biol.* 14, 99–114. doi: 10.1186/1471-2229-14-99

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456 (7221), 470–476. doi: 10.1038/nature07509

Wang, H., You, C., Chang, F., Wang, Y., Wang, L., Qi, J., et al. (2014). Alternative splicing during Arabidopsis flower development results in constitutive and stage-regulated isoforms. *Front. Genet.* 5, 25. doi: 10.3389/fgene.2014.00025

Wang, T., Wang, H., Cai, D., Gao, Y., Zhang, H., Wang, Y., et al. (2017). Comprehensive profiling of rhizome-associated alternative splicing and alternative polyadenylation in moso bamboo (*Phyllostachys edulis*). *Plant J.* 91 (4), 684–699. doi: 10.1111/tpj.13597

Wang, M., Wang, P., Liang, F., Ye, Z., Li, J., Shen, C., et al. (2018). A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *New Phytol.* 217 (1), 163–178. doi: 10.1111/nph.14762

Wang, J.-l., Zhang, Y., Pan, X.-D., Du, J.-J., Ma, L.-M., and Guo, X.-Y. (2019). Discovery of leaf region and time point related modules and genes in maize (Zea mays L.) leaves by Weighted Gene Co-expression Network analysis (WGCNA) of gene expression profiles of carbon metabolism. *J. Integr. Agric.* 18 (2), 350–360. doi: 10.1016/s2095-3119(18)62029-5

Wang, P., Dai, L., Ai, J., Wang, Y., and Ren, F. (2019). Identification and functional prediction of cold-related long non-coding RNA (lncRNA) in grapevine. *Sci. Rep.* 9 (1), 6638. doi: 10.1038/s41598-019-43269-5

Wang, Z., Zhang, H., and Gong, W. (2019). Genome-wide identification and comparative analysis of alternative splicing across four legume species. *Planta* 249 (4), 1133–1142. doi: 10.1007/s00425-018-03073-3

Workman, R. E., Myrka, A. M., Wong, G. W., Tseng, E., Welch, K. C.Jr., and Timp, W. (2018). Single-molecule, full-length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird Archilochus colubris. *Gigascience* 7 (3), 1–12. doi: 10.1093/gigascience/giy009

Wu, T. D., and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21 (9), 1859–1875. doi: 10.1093/bioinformatics/bti310

Xu, Q., Modrek, B., and Lee, C. (2002). Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res.* 30 (17), 3754–3766. doi: 10.1093/nar/gkf492

Xu, X., Zhang, H., Xie, T., Xu, Y., Zhao, L., and Tian, W. (2017). Effects of Climate Change on the Potentially Suitable Climatic Geographical Range of *Liriodendron chinense*. *Forests* 8 (10), 399–412. doi: 10.3390/f8100399

Yan, X. B., Tang, C. H., Huang, Y., Fang, H., Yu, Z. Q., Wu, L. M., et al. (2010). Alternative splicing in exon 9 of glucocorticoid receptor pre-mRNA is

regulated by SRp40. *Mol. Biol. Rep.* 37 (3), 1427–1433. doi: 10.1007/s11033-009-9529-z

Yang, Y., Xu, M., Luo, Q., Wang, J., and Li, H. (2014). De novo transcriptome analysis of *Liriodendron chinense* petals and leaves by Illumina sequencing. *Gene* 534 (2), 155–162. doi: 10.1016/j.gene.2013.10.073

Yang, Z., Yang, Z., Yang, C., Wang, Z., Chen, D., Xie, Y., et al. (2019). Identification and genetic analysis of alternative splicing of long non-coding RNAs in tomato initial flowering stage. *Genomics* 112 (1), 897–907. doi: 10.1016/j.ygeno.2019.06.005

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11, R14–R25. doi: 10.1186/gb-2010-11-2-r14

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, Article17. doi: 10.2202/1544-6115.1128

Zhang, H., Mao, R., Wang, Y., Zhang, L., Wang, C., Lv, S., et al. (2019). Transcriptome-wide alternative splicing modulation during plant-pathogen interactions in wheat. *Plant Sci.* 288, 110160. doi: 10.1016/j.plantsci.2019.05.023

Zhang, Y., Nyong, A. T., Shi, T., and Yang, P. (2019). The complexity of alternative splicing and landscape of tissue-specific expression in lotus (*Nelumbo nucifera*)

unveiled by Illumina- and single-molecule real-time-based RNA-sequencing. *DNA Res.* 26 (4), 301–311. doi: 10.1093/dnares/dsz010

Zhao, H., Gao, Z., Wang, L., Wang, J., Wang, S., Fei, B., et al. (2018). Chromosome-level reference genome and alternative splicing atlas of moso bamboo (*Phyllostachys edulis*). *Gigascience* 7 (10), 1–12. doi: 10.1093/gigascience/giy115

Zhu, M., Xie, H., Wei, X., Dossa, K., Yu, Y., Hui, S., et al. (2019). WGCNA Analysis of Salt-Responsive Core Transcriptome Identifies Novel Hub Genes in Rice. *Genes (Basel)* 10 (9), 719–734. doi: 10.3390/genes10090719