



Article

Comprehensive Profiling of Alternative Splicing and Alternative Polyadenylation during Fruit Ripening in Watermelon (*Citrullus lanatus*)

Yongtao Yu ^{1,†}, Yuxiang Liufu ^{2,†}, Yi Ren ¹, Jie Zhang ¹, Maoying Li ¹ , Shouwei Tian ¹ , Jinfang Wang ¹, Shengjin Liao ¹, Guoyi Gong ¹, Haiying Zhang ^{1,*} and Shaogui Guo ^{1,*}

¹ National Watermelon and Melon Improvement Center, Beijing Academy of Agriculture and Forestry Sciences, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), National Engineering Research Center for Vegetables, Beijing Key Laboratory of Vegetable Germplasm Improvement, Beijing 100097, China; wangjinfang@nercv.org (J.W.)

² College of Forestry, Basic Forestry and Proteomics Research Center, Fujian Agriculture and Forestry University, Fuzhou 350002, China

* Correspondence: zhanghaiying@nercv.org (H.Z.); guoshaogui@nercv.org (S.G.)

† These authors contributed equally to this manuscript.

Abstract: Fruit ripening is a highly complicated process that is accompanied by the formation of fruit quality. In recent years, a series of studies have demonstrated post-transcriptional control play important roles in fruit ripening and fruit quality formation. Till now, the post-transcriptional mechanisms for watermelon fruit ripening have not been comprehensively studied. In this study, we conducted PacBio single-molecule long-read sequencing to identify genome-wide alternative splicing (AS), alternative polyadenylation (APA) and long non-coding RNAs (lncRNAs) in watermelon fruit. In total, 6,921,295 error-corrected and mapped full-length non-chimeric (FLNC) reads were obtained. Notably, more than 42,285 distinct splicing isoforms were derived from 5,891,183 intron-containing full-length FLNC reads, including a large number of AS events associated with fruit ripening. In addition, we characterized 21,506 polyadenylation sites from 11,611 genes, 8703 of which have APA sites. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis showed that fructose and mannose metabolism, starch and sucrose metabolism and carotenoid biosynthesis were both enriched in genes undergoing AS and APA. These results suggest that post-transcriptional regulation might potentially have a key role in regulation of fruit ripening in watermelon. Taken together, our comprehensive PacBio long-read sequencing results offer a valuable resource for watermelon research, and provide new insights into the molecular mechanisms underlying the complex regulatory networks of watermelon fruit ripening.

Keywords: watermelon; alternative splicing; alternative polyadenylation; lncRNA; fruit ripening



Citation: Yu, Y.; Liufu, Y.; Ren, Y.; Zhang, J.; Li, M.; Tian, S.; Wang, J.; Liao, S.; Gong, G.; Zhang, H.; et al. Comprehensive Profiling of Alternative Splicing and Alternative Polyadenylation during Fruit Ripening in Watermelon (*Citrullus lanatus*). *Int. J. Mol. Sci.* **2023**, *24*, 15333. <https://doi.org/10.3390/ijms242015333>

Academic Editor: María Serrano

Received: 30 August 2023

Revised: 5 October 2023

Accepted: 7 October 2023

Published: 18 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Watermelon (*Citrullus lanatus*, $2n = 2 \times = 22$), a typical non-climacteric fruit, originated in Africa, belongs to the xerophytic genus *Citrullus* Schrad. ex Eckl. et Zeyh. of the botanical family *Cucurbitaceae* and is an important cucurbit crop grown throughout the world. Watermelon is an important experimental model to study the molecular mechanisms of non-climacteric fruits' ripening [1]. This reflects its economic value and many favorable genetic characteristics, such as relatively small genome size (359.8 Mb), diploid cultivar, short life cycle (3–4 months), and efficient genetic transformation system [2].

Ripening of fleshy fruits is a highly complex process that involves dramatic changes in sugar content, fruit color, fruit texture, flavor, and aroma [3,4]. In recent years, the mechanisms of fruit ripening have been studied extensively at the transcriptional level, and many transcription factors as well as 5-methylcytosine (5mC) DNA methylation at various gene loci have been identified and proven to be responsible for transcriptional regulation

of gene expression during the ripening process [5,6]. In addition, a series of studies have demonstrated that posttranscriptional control also plays vital roles in fruit ripening such as AS, APA, mRNA N6-methyladenosine RNA modification methylation and noncoding RNAs [7–9].

AS and APA are two important post-transcriptional regulatory mechanisms that increase transcriptomic and proteomic diversity in eukaryotic organisms [10,11]. AS is the process through which multiple transcript isoforms are produced from a single gene [12]. Previous short-read RNA-seq analysis has shown that up to 40–60% of intron-containing genes are alternatively spliced in different tissues or developmental stages, suggesting AS plays key roles in the regulation of development in the model plants, such as *Arabidopsis thaliana* and rice [13–16]. APA is a widespread phenomenon, generating mRNAs with distinct 3' ends [11]. The large-scale studies using transcriptome-wide techniques have revealed that 70% of *Arabidopsis* genes and around 50% of rice genes have at least one alternative poly(A) site [17,18]. Single-molecule real-time sequencing (SMRT), developed by Pacific BioSciences (PacBio), offers longer read lengths than the RNA-seq which are needed for transcriptome assembly, making it well-suited for unsolved problems in transcriptome research [19].

Over the past decade, single molecule sequencing technology has been used to characterize the transcriptome diversity in *Salvia miltiorrhiza* [20], *Z. mays* [21], *Sorghum bicolor* [22], *Fragaria vesca* [23], *Phyllostachys edulis* [24], *Trifolium pratense* L. [25], *Ricinus communis* [26], *Ginkgo biloba* [27], *Populus* [28], *Solanum betaceum Cav* [29] and so on. In these studies, most of the full-length cDNAs were generated by combining in equal amounts total RNA from different tissues and/or did not make replicates. In this study, we firstly carried out SMRT long-read sequencing of four critical developmental stages in watermelon (three repeats for each stage). A total of 7,837,450 reads of inserts (ROIs) were generated for 12 samples. Finally, 6,921,295 error-corrected and mapped FLNC reads were obtained. The inclusion of more high-quality FLNC reads is more conducive to conduct AS and APA analyses, which can better reflect the dynamic changes during fruit ripening at post-transcriptional level.

2. Results

2.1. Watermelon Fruit Transcriptome Sequencing Using PacBio Iso-Seq

Samples were collected at four key developmental stages (10 days after pollination (DAP), 18 DAP, 26 DAP and 34 DAP) (Figure 1), and each stage had three repeats. These samples were the same as in our previously published study [30]. In this current study, PacBio sequel II platform was used to generate long reads for the above 12 samples. A total of 7,837,450 reads of inserts (ROIs) were generated for 12 samples, and on average 88% of them were full-length non-chimeric (FLNC) reads with the entire transcript region from the 5' to the 3' end based on the inclusion of barcoded primers and the 3' poly(A) tails. A total of 843,963,605 paired-end (PE) reads from our previously published transcriptomic data were used to correct the FLNC reads, then the corrected FLNC reads were mapped to the watermelon genome (97103, V2.5 version) using minimap2 [31]. Finally, we obtained 6,921,295 error-corrected and mapped (high-quality) FLNC reads, and the average FLNC reads of 12 samples is 576,774 (Table 1). These FLNC reads cover 73.49% of annotated genes, and 14,520 genes were supported by at least three PacBio reads (Supplementary Figure S1). The repeatability analysis of 12 samples showed that the results of APA quantification were reliable (Supplementary Figure S2).

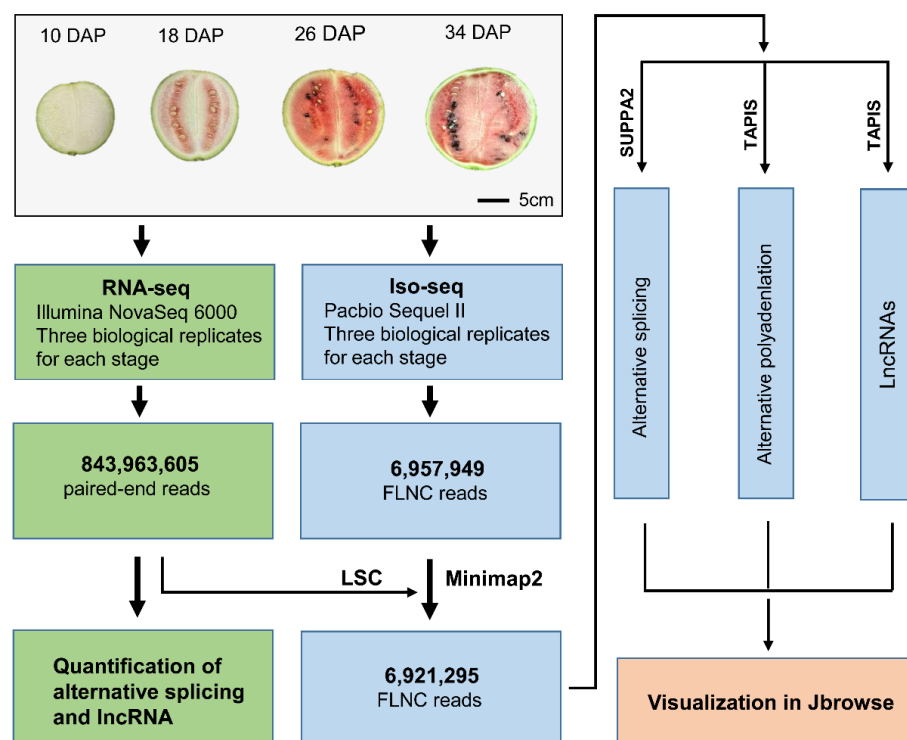


Figure 1. Flowchart of the experimental design and analysis for PacBio long-read sequencing and RNA-seq. In total, 6,921,295 mapped full-length non-chimeric (FLNC) reads were used for the downstream analysis, including alternative splicing (AS), alternative polyadenylation (APA) and lncRNA.

Table 1. Summary of PacBio sequencing.

Sample	cDNAsize	Number of Reads of Insert	Number of Five-Five Primer Reads	Number of Three-Three Primer Reads	Number of Adaptor Reads	Number of Filtered Short Reads	Number of Full-Length Non-Chimeric Reads	Error-Corrected and Mapped FLNC
DAP10-1	1-10k	632,678	19,259	7905	3	0	580,464	577,193
DAP10-2	1-10k	726,681	12,245	22,527	12	0	663,057	661,727
DAP10-3	1-10k	639,781	146,488	7345	18	0	463,837	460,828
DAP18-1	1-10k	727,467	20,973	7599	14	0	681,511	679,326
DAP18-2	1-10k	696,383	104,864	12,243	9	0	557,964	555,213
DAP18-3	1-10k	690,537	29,087	11,170	15	0	629,609	626,298
DAP26-1	1-10k	618,696	21,428	12,025	6	0	573,505	569,181
DAP26-2	1-10k	707,380	24,627	14,069	7	0	654,125	651,197
DAP26-3	1-10k	550,489	13,507	5581	7	0	520,269	516,792
DAP34-1	1-10k	587,052	98,357	18,328	42	0	457,391	453,471
DAP34-2	1-10k	531,201	11,605	5697	3	0	502,478	499,375
DAP34-3	1-10k	729,105	27,788	14,078	3	0	673,739	670,694

2.2. Identification of Full-Length Splicing Isoforms

Till now, the study of the watermelon transcriptome is not totally complete, in particular the annotation of full-length splice isoforms. We obtained a total of 5,891,183 intron-containing FLNC reads that collapsed into 42,285 distinct full-length splice isoforms representing 135,772 unique introns. The consensus donor sites and acceptor sites of watermelon are TTAAAG|GTAAG and TGCAG|GTAAA, respectively (Figure 2A). The consensus donor and acceptor sites were similar to those in rice and moso bamboo [24,32]. The median number of introns in intron-containing genes in Arabidopsis [33] and moso bamboo [24] is four and seven, respectively, whereas that for each unique splicing isoform in watermelon is six. Due to the limit of read length, it is challenging to identify full-length splicing isoforms accurately by RNA-seq, while PacBio long reads could produce more

comprehensive annotation of splicing isoforms with superior contiguity [24,34]. Therefore, we examined the full-length transcript structures of distinct splicing isoforms using PacBio long-read sequencing data. In total, 16,495 AS events were identified without transcriptome assembly, so any possible artificial result could be avoided (Supplementary Data S1). We further classified these AS events into seven distinct types: 5574 retained introns (RI), 3557 skipping exon (SE), 3104 alternative 5' splice-site (A5), 3141 alternative 3' splice-site (A3), 697 alternative first exon (AF), 267 alternative last exon (AL) events and 155 mutually exclusive exons (MX) (Supplementary Figure S3A, Supplementary Data S1). KEGG analysis of genes undergoing AS showed that fructose and mannose metabolism, starch and sucrose metabolism and carotenoid biosynthesis were enriched, which included *tonoplast sugar transporter 2 (TST2)*, *Sugars Will Eventually Be Exported Transporter 3 (SWEET3)*, *insoluble acid invertase (IAI, Cla97C05G099220)*, *UDP-Gal/Glc PPase (UGGP, Cla97C08G148810)*, *Phytoene synthase (PSY, Cla97C01G008760)* and *Phytoene desaturase (PDS, Cla97C07G142100)* (Supplementary Figure S4, Supplementary Data S2 and S3). All the enriched KEGG terms are listed in Supplementary Data S2. Next, we validated the retained intron event of *PDS (Cla97C02G037750, carotenoid biosynthetic enzyme)* by RT-PCR. Consistently, RT-PCR could amplify two bands, which also meet the expected size (Figure 2B,C). Considering the relatively low sequencing depth of PacBio long-read sequencing, we quantified differential splicing events using Illumina short reads. The results are shown in Supplementary Figure S3B–D and Supplementary Data S3.

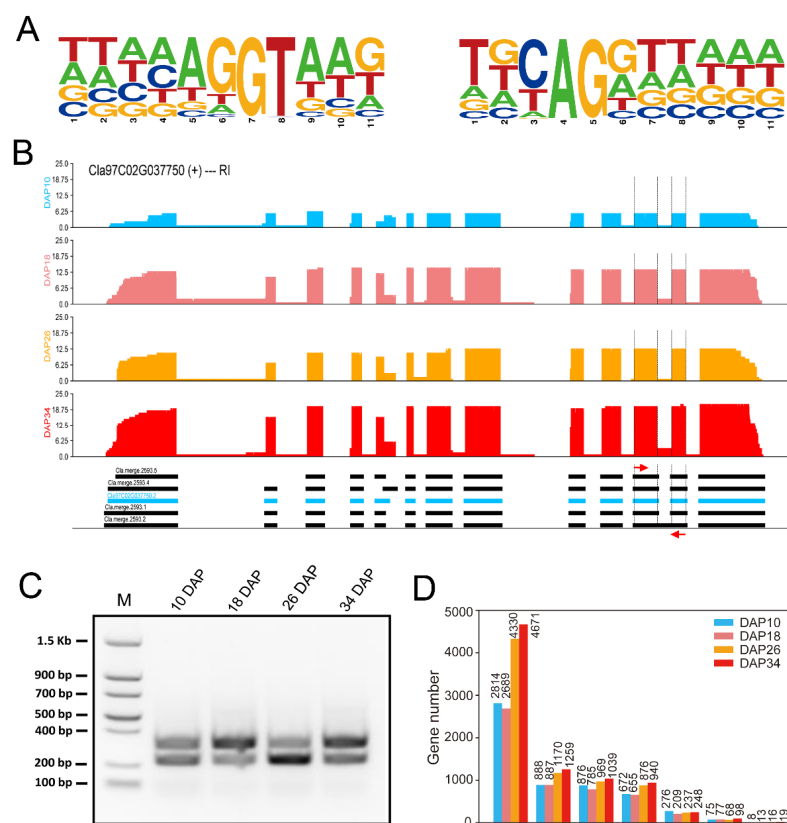


Figure 2. Comprehensive identification of splicing isoforms in watermelon fruit. (A) The nucleotide distributions flanking the 5' donor (left) or 3' acceptor sites (right). (B) Wiggle plot showing the intron retention event of the *PDS* gene (Cla97C02G037750). The red arrows represent the primers' position. (C) Validation of the intron retention event in Figure 3B using RT-PCR. (D) Statistics of total AS events in watermelon fruit. RI: retained introns; SE: skipping exon; A5: alternative 5' splice-site; A3: alternative 3' splice-site; AF: alternative first exon; AL: alternative last exon; MX: mutually exclusive exons (MX).

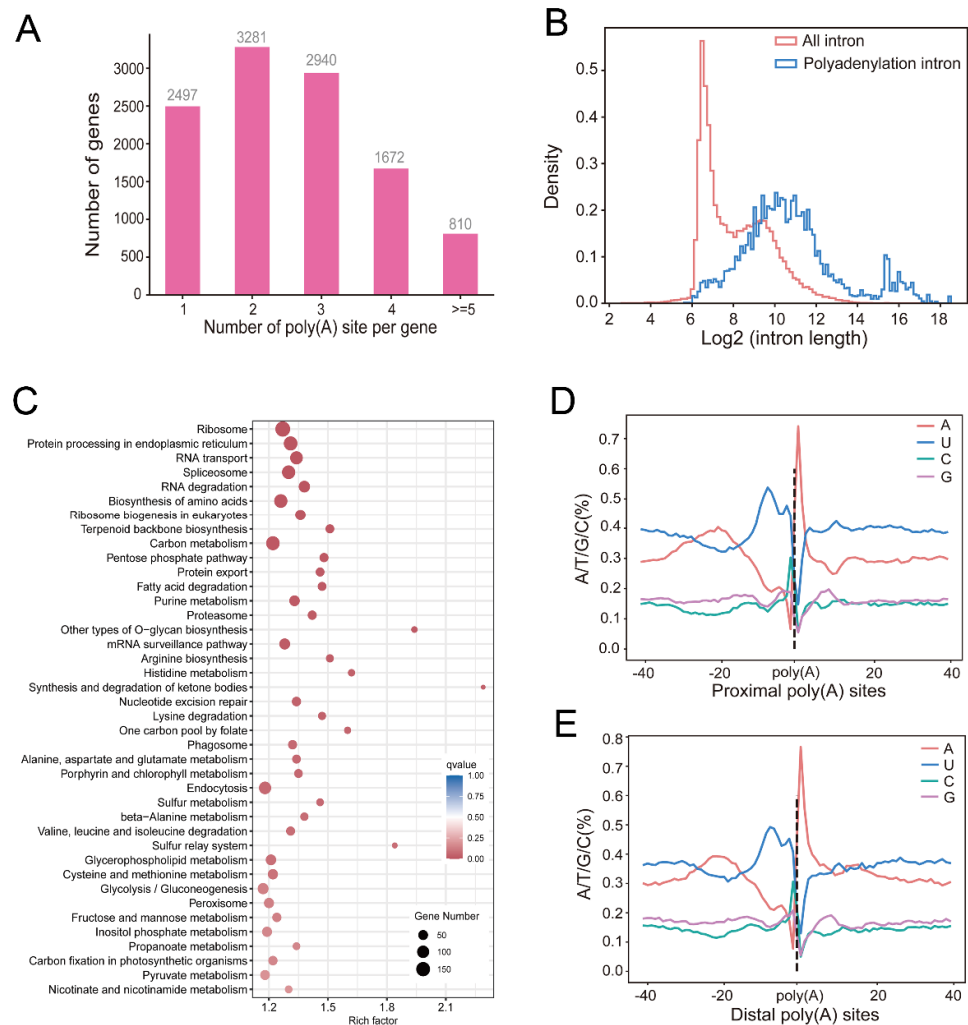


Figure 3. Genome-wide identification of alternative polyadenylation in watermelon fruit. **(A)** Distribution of the number of poly(A) sites per gene. **(B)** Length distribution introns with polyadenylation cleavage sites and all introns. **(C)** KEGG analysis of genes undergoing APA. **(D,E)** The nucleotide composition profile around proximal and distal poly(A) sites, respectively.

2.3. Profiling of Global APA Sites and Differential APA

Global establishment of the accurate polyadenylation cleavage sites to define the 3' end of genes will greatly improve the gene annotation. PacBio long-read sequencing produces full-length of cDNA, which is suitable to identify APA events. In this study, we firstly carried out long-read sequencing to study the global polyadenylation events at single-nucleotide resolution in watermelon. In total, we characterized 21,506 polyadenylation sites from 11,611 genes, and 8703 genes contained two or more polyadenylation sites (Figure 3A, Supplementary Data S4). Furthermore, a total of 4959 intronic polyadenylation sites were identified (Supplementary Data S4). Interestingly, these introns were longer than those without polyadenylation sites (Figure 3B). Next, we validated the APA event of *Cla97C01G008870* (Isocitrate dehydrogenase (ICDH), citric acid biosynthetic enzyme) by 3'-RACE. Consistently, PCR could amplify multiple bands, which represented different 3' ends for transcript variants (Supplementary Figure S5). Motifs around the watermelon polyadenylation region contained UGUA and AAUAAA (Supplementary Figure S6A), which were similar to previous studies in *Sorghum bicolor*, *Ginkgo biloba* and moso bamboo [22,24,27]. Proximal and distal poly(A) cleavage sites displayed extremely similar distribution, suggesting that proximal sites are real poly(A) cleavage sites (Figure 3D–E). Similar to AS, KEGG analysis showed that fructose and

mannose metabolism, starch and sucrose metabolism and carotenoid biosynthesis were also enriched in genes undergoing APA (Figure 3C, Supplementary Data S5), which included *IAI* (*Cla97C05G099220*), *Sucrose Synthase* (*SUS*, *Cla97C10G194010*), *Trehalose 6-phosphate phosphatase* (*TPP*, *Cla97C07G136350*, *Cla97C01G018360*), *PSY* (*Cla97C01G008760*, *Cla97C07G137500*), *PDS* (*Cla97C07G142100*, *Cla97C10G198770*) and *9-cis-epoxy-carotenoid dioxygenase* (*NCED*, *Cla97C07G137260*, *Cla97C01G024630*). All the enriched KEGG terms are listed in Supplementary Data S5. Finally, we quantified differential APA, and the results are shown in Supplementary Figure S6B–D and Supplementary Data S6.

2.4. Identification of Long Non-Coding RNA from PacBio Sequences

In plants, lncRNAs are largely involved in regulation of plant growth and development, synthesizing secondary metabolites, and responding to environmental stress [35]. Recently, a series of noncoding RNAs (ncRNAs), especially lncRNAs, have been functionally characterized in fruit quality formation and fruit ripening at the posttranscriptional level [7]. Using PacBio long-read data, we identified a total of 594 high-confidence lncRNAs in watermelon fruit, which presented a shorter exon length than that of coding genes (Figure 4A, Supplementary Data S7). Furthermore, 72 lncRNAs were up-regulated and 79 lncRNAs were down-regulated in 10 DAP vs. 18 DAP group, 37 lncRNAs were up-regulated and 12 lncRNAs were down-regulated in 18 DAP vs. 26 DAP group, and 26 lncRNAs were up-regulated and 14 lncRNAs were down-regulated in 26 DAP vs. 34 DAP group (Figure 4B–D, Supplementary Data S8). These differentially expressed lncRNAs might play important roles in watermelon fruit ripening.

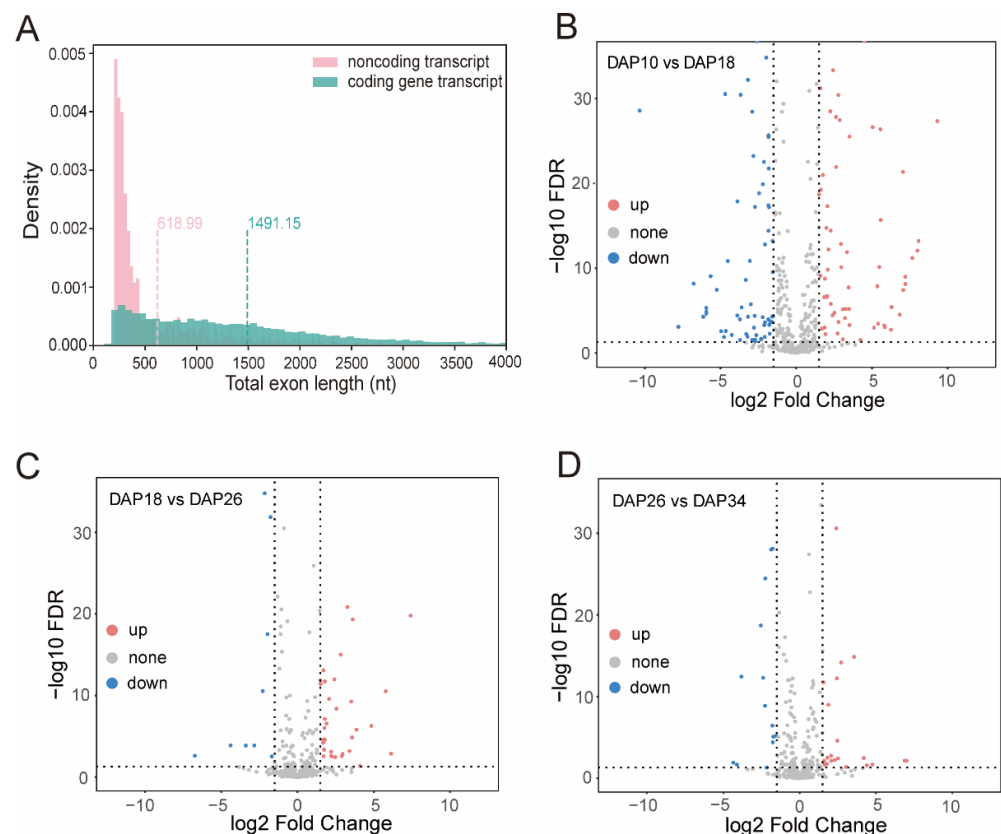


Figure 4. Identification of lncRNA in watermelon fruit. (A) The density distribution of exon length from lncRNA transcripts and protein-coding transcripts. Volcano plot showing differentially expressed lncRNAs: (B) in DAP10 vs. DAP18 group; (C) in DAP18 vs. DAP26 group; and (D) in DAP26 vs. DAP34 group.

3. Discussion

Watermelon is an important fruit crop, and an ideal model plant for studies on non-climacteric fruit ripening. In this study, using PacBio single-molecule long-read sequencing, we present the full-length transcriptome of four key developmental stages in watermelon. In total, we obtained 6,921,295 error-corrected and mapped FLNC reads. In addition, 843,963,605 paired-end (PE) reads from our previously published study were also used in this work. These high-quality datasets provided a solid foundation for our subsequent analyses.

AS and APA play a crucial role in the functional diversity of a gene by generating different isoforms to regulate gene expression at the post-transcriptional level in higher eukaryotes. To date, genome-wide analyses of alternative splicing have been studied in some fruits, including tomato [36], kiwifruit [37], olive (*Olea europaea*) [38], papaya [39], peach [39], melon [39], cucumber [39], blueberry [40], citrus [41], grape (*Vitis vinifera*) [42], sweet cherry (*Prunus avium*) [43], strawberry [23,44] and watermelon (*Citrullus lanatus*) [45]. All the studies conducted RNA-seq to identify AS events except in strawberry and olive, which employed PacBio long-read sequencing. In cucumber and melon, AS was significantly more prevalent at the ripe stage than at the immature stage, while the opposite trend was shown in papaya and peach, implying that developmental stages adopt different alternative splicing strategies for their specific functions [39]. In addition, all vegetative tissues of tetraploid watermelon except the fruit displayed an increased level of AS than diploid watermelon throughout the growth period [45]. In this study, we found AS events were significantly increased at the ripe stage (26 DAP and 34 DAP) than at the immature stage (Figure 2D), which was similar to cucumber and melon.

Common RNA-seq could not provide a better understanding of genome-wide profiling of polyadenylation sites [46]. Previous developed polyadenylation site sequencing (PAS-seq) was used to comprehensively identify plant APA events [17,47]. However, the method has an internal priming issue in that the oligo (dT) primer could also bind internal A-rich sequences [48]. PacBio long-read sequencing generated full-length structures of each splicing isoform and allowed the identification of exact polyadenylation sites [22]. Till now, global APA identification during fleshy fruit development has rarely been reported. In this study, PacBio sequencing was employed to generate long reference sequences and profiling of APA from 8703 genes in watermelon. KEGG analysis showed that fructose and mannose metabolism, starch and sucrose metabolism and carotenoid biosynthesis were enriched in genes undergoing AS and APA, respectively. Interestingly, *IAI* (*Clu97C05G099220*), (*PSY*, *Clu97C01G008760*) and *PDS* (*Clu97C07G142100*) were simultaneously regulated by both AS and APA. Taken together, these results suggested that AS and APA might play important roles in watermelon fruit ripening. In total, 4959 introns from 3073 genes included intronic polyadenylation, which might encode shorter proteins (Ozso-lak et al., 2010). In *Arabidopsis* and moso bamboo, introns with polyadenylation sites always contain transposable elements (TEs), which result in introns with polyadenylation sites are longer than that without polyadenylation sites [24,49,50]. In addition, previous studies showed that association of intronic TE elements through epigenetic regulation (DNA methylation, histone modification) impact the intronic polyadenylation in *Arabidopsis* [49,50]. In this study, we also found that introns with polyadenylation sites are longer in watermelon fruit (Figure 3B). In future, more fruit ripening-related epigenetic data will reveal the interplay between epigenetic regulation and intronic polyadenylation in watermelon, such as histone modification and DNA methylation.

Based on our previously published transcriptomic data, we found that a few splicing-related factors displayed obvious up-regulated expression during watermelon fruit ripening such as serine/arginine-rich splicing factor (SR), DEAD-box ATP-dependent RNA helicase (DDX), cyclin-dependent kinase g-2 (CDK), and pre-mRNA-processing factor 39 (PRP39) (Figure 5). Additionally, APA-associated factors also showed different expression during ripening such as cleavage and polyadenylation specificity factor (CPSF), poly(A) polymerase (PAP) and pre-mRNA cleavage factor Im 25 kDa subunit 1 (CFIm25) (Figure 5).

These differential splicing-related and APA-associated genes might take part in regulation of AS and APA during fruit ripening in watermelon.

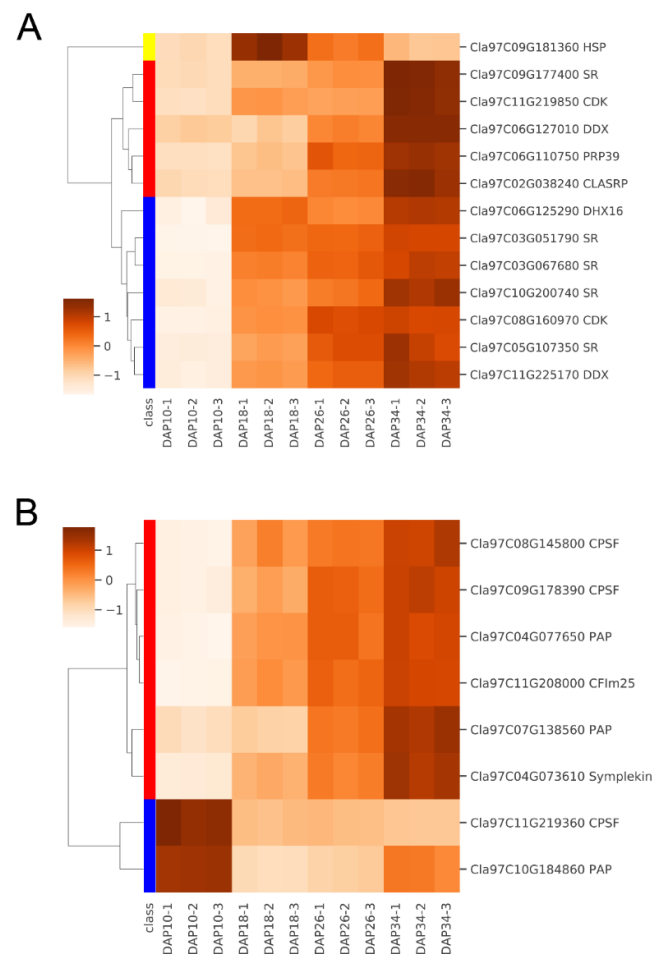


Figure 5. Heatmap of splicing factors and poly(A) factors in watermelon fruit. **(A)** Splicing-related factors. SR: serine/arginine-rich splicing factor; DDX: DEAD-box ATP-dependent RNA helicase; CDK: cyclin-dependent kinase; PRP39: pre-mRNA-processing factor 39; DHX16: DEAH-box protein, ATP-dependent RNA helicase; HSP: 35S_U5_associated_proteins; CLASRP: CLK4-associating serine/arginine rich protein. **(B)** APA-related factors. CPSF: cleavage and polyadenylation specificity factor; PAP: poly(A) polymerase; CFIm25: pre-mRNA cleavage factor Im 25 kDa subunit.

Climate change will have profound effects on food production. The production and quality of fresh fruit can be directly or indirectly affected by exposure to high temperatures and elevated levels of carbon dioxide and ozone. Accelerated crop improvement (traditional breeding methods, genetic modification and gene editing), and new cultivation and processing techniques are required to produce sufficient food supplies and/or economize the production cost, so as to meet the basic nutrient requirements of the growing human population in this age of climate change. Our comprehensive PacBio long-read sequencing present a comprehensive view of fruit ripening-related AS, APA and lncRNA, which provides a better understanding of the complex post-transcriptional regulatory networks of fruit ripening in watermelon. Knowledge of the gene function and regulatory networks can make a significant contribution to the development of modern varieties that are able to withstand and adapt to the climatic conditions brought about by climate change.

4. Materials and Methods

4.1. Plant Materials

Watermelon *C. lanatus* (Thunb.) Matsum. & Nakai subsp. *vulgaris* cv 97,103 was used in this study. In order to make the fruit grow uniformly, we kept only one watermelon per plant. Flowers were hand-pollinated at 3–5 nodes and tagged. PacBio long-read sequencing was performed on three biological replicates of watermelon center flesh samples; these were collected at four developmental stages (10, 18, 26 and 34 DAP (days after pollination), Figure 1), each replicate resulting from the pooling of at least 10 fruits. Samples were rapidly frozen in liquid nitrogen and stored at -80°C . All the samples used in this study were the same as from our previous study [30].

4.2. Total RNA Extraction, Construction of PacBio Library and Sequencing

Total RNA was isolated with a Total RNA Rapid Extraction Kit (Huayueyang Biotechnologies Co. Ltd., Beijing, China). Three biological replicates were carried out for Iso-Seq. The purity and concentration of total RNA were assessed using NanoDrop 2000, and the RNA integrity was evaluated with the Agilent 2100/4200 system. A total of 2 μg RNA was used for library construction. PacBio single-molecule real-time (SMRT) bell library preparation was performed using SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, PN 101-853-100) in accordance with the manufacturer's instructions. In brief, first strand cDNA was synthesized by reverse transcription after adding primer with Oligo dT paired with poly(A). Template switching was performed by adding the other primer paired with TS oligo. Full-length cDNA was obtained by cDNA amplification with barcoded cDNA PCR primers. Amplified cDNA was purified ProNex Beads and quantified by Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). SMRTbell library was prepared using the Kit 2.0. Single-strand overhangs were removed, and then we performed DNA damage repair, end-repair, A-tailing and adapter ligation. After library purification using 0.45X AMPure[®] PB Beads and 1–10 Kb size selection, the library was prepared for sequencing for 30 h movie on the Sequel II system (Pacific Biosciences, Menlo Park, CA, USA).

4.3. Data Analysis

Watermelon reference genome and annotated gene models (97103, V2.5 version) were downloaded from CuGenDB (<http://cucurbitgenomics.org/>, accessed on 9 October 2023). The generated raw data were processed using the SMRT Analysis software (ISOseq version3.0). HiFi reads were generated by SMRTLink 10.0 software with parameters `--min-passes = 3 --min-rq = 0.99`. According to whether the sequence contained 5' primers, 3' primer or the poly(A) tail, the HiFi reads were divided into full-length and non-full-length sequences. Only the HiFi reads without any inter primer were considered as a full-length non-chimeric (FLNC) read. The FLNC were corrected by RNA-seq reads via LSC 1.alpha with Bowtie 2 v.2.2.1 alignment [51]. The error-corrected FLNC reads were then mapped to watermelon reference genome using minimap2 by the command `(minimap2 -ax splice uf k14--sam-hit-only secondary=no)` [31]. Alignment files were sorted, indexed, and filtered (`-F 2048`) by SAMtools. Stringtie2 (V2.2.1) was used for transcript assembly with the option `-L --conservative` [52]. Next, the collapsed transcripts were compared to reference annotation by Gffcompare (V0.12.6) with default parameter. MEME (V5.5.3) was used to search the upstream and downstream cis elements of polyadenylation sites [53]. KEGG pathway enrichment analyses were performed by the ClusterProfiler R- package, and the ggplot2 R-package was used for visualization [54].

4.4. Quantification of Differential Alternative Splicing

The AS event was identified by SUPPA2 (V2.3) generateEvents with options `'-e SE SS MX RI FL -f ioe'` [55]. The quantification of AS event was performed by Salmon (V1.10.1) with options `'--gcBias --validateMappings'` using RNA-seq. The differential alternative splicing events were quantified by SUPPA2 diffSplice using the option `--method (empirical)`. Throughout, AS events with $p < 0.05$ were regarded as differential AS events.

4.5. Quantification of Differential Alternative Polyadenylation

Poly(A) site with >5% usage rate was remained as candidate poly(A) site. The distance between two distinct poly(A) sites should be more than 30 nt. The reads with termination site located 12 nt immediately upstream and downstream a poly(A) site was assigned to the poly(A) site. The poly(A) site with higher abundance was selected as the major sites. The abundance for each poly(A) site was normalized based on read count from the highest sites and library size. Fisher's exact test was used for differential APA sites among the DAP10 vs. DAP18 group, DAP18 vs. DAP26 group and DAP26 vs. DAP34 group ($p < 0.05$).

4.6. Identification of lncRNA

For lncRNA identification, FLNC reads were merged together and aligned to genome by minimap2 [31] to generate BAM files, which were assembled to the transcripts by Stringtie2 [52]. Transcripts with only one exon and exon lengths of <200 bp were filtered out from downstream analysis. The prediction of lncRNA were preformed by CPC2, PLEK, lncFinder and Pfam with default options. The expression of lncRNA was quantificatied by Salmon with RNA-seq data.

4.7. PCR Validation of AS and APA

Two micrograms of total RNA was subjected to first-strand cDNA synthesis using a TransScript First-Strand cDNA Synthesis SuperMix (TransGen Biotech, Cat# AT301-02, Beijing, China) and an oligo (dT18) primer. AS was validated by RT-PCR using $2 \times$ Phanta Max Master Mix (Dye Plus) (Vazyme, Cat# P525-01, Nanjing, China) with $5 \times$ diluted cDNA as the template. The PCR products were visualized in 1.2% agarose gel stained by GelStain. The APA events were validated using a 3' RACE Kit (Clontech, cat # 6106) according to the manufacturer's instructions. Primer lists used in the validation are listed in Supplementary Table S1.

4.8. Data Presentation

All the statistical analyses and plots were carried out using Python matplotlib packages (version 3.8.0).

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms242015333/s1>.

Author Contributions: Y.Y., Y.L., H.Z. and S.G. designed the study and wrote the manuscript with contributions from all coauthors. S.L. provided good advice for the data analysis. Y.R., J.Z. and M.L. collected and grew the plant material. S.T., J.W. and G.G. provided critical comments and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Natural Science Foundation of China (NSFC Grant No. 32172603), Young Beijing Scholar Program (YBSP019), Foundation of Beijing Academy of Agricultural and Forestry Sciences (QNJJ202334), Construction of Cucurbits Collaboration and Innovation Center (XTCX202301) and Innovation and Development Program of Beijing Vegetable Research Center (KYCX202301).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Our PacBio Iso-seq data were submitted to the Sequence Read Archive (SRA) of NCBI under accession number PRJNA718122.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Erickson, D.L.; Smith, B.D.; Clarke, A.C.; Sandweiss, D.H.; Tuross, N. An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18315–18320. [[CrossRef](#)]
2. Guo, S.; Zhang, J.; Sun, H.; Salse, J.; Lucas, W.J.; Zhang, H.; Zheng, Y.; Mao, L.; Ren, Y.; Wang, Z.; et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* **2013**, *45*, 51–58. [[CrossRef](#)]

3. Giovannoni, J. Molecular Biology of Fruit Maturation and Ripening. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 725–749. [[CrossRef](#)]
4. Seymour, G.B.; Østergaard, L.; Chapman, N.H.; Knapp, S.; Martin, C. Fruit development and ripening. *Annu. Rev. Plant Biol.* **2013**, *64*, 219–241. [[CrossRef](#)]
5. Chen, T.; Qin, G.; Tian, S. Regulatory network of fruit ripening: Current understanding and future challenges. *New Phytol.* **2020**, *228*, 1219–1226. [[CrossRef](#)]
6. Sánchez-Gómez, C.; Posé, D.; Martín-Pizarro, C. Insights into transcription factors controlling strawberry fruit development and ripening. *Front. Plant Sci.* **2022**, *13*, 1022369. [[CrossRef](#)]
7. Wang, W.; Wang, Y.; Chen, T.; Qin, G.; Tian, S. Current insights into posttranscriptional regulation of fleshy fruit ripening. *Plant Physiol.* **2023**, *192*, 1785–1798. [[CrossRef](#)]
8. Tang, D.; Gallusci, P.; Lang, Z. Fruit development and epigenetic modifications. *New Phytol.* **2020**, *228*, 839–844. [[CrossRef](#)]
9. Ji, Y.; Wang, A. Recent advances in epigenetic triggering of climacteric fruit ripening. *Plant Physiol.* **2023**, *192*, 1711–1717. [[CrossRef](#)]
10. Kalsotra, A.; Cooper, T.A. Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.* **2011**, *12*, 715–729. [[CrossRef](#)]
11. Elkon, R.; Ugalde, A.P.; Agami, R. Alternative cleavage and polyadenylation: Extent, regulation and function. *Nat. Rev. Genet.* **2013**, *14*, 496–506. [[CrossRef](#)]
12. Nilsen, T.W.; Graveley, B.R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **2010**, *463*, 457–463. [[CrossRef](#)]
13. Filichkin, S.A.; Priest, H.D.; Givan, S.A.; Shen, R.; Bryant, D.W.; Fox, S.E.; Wong, W.-K.; Mockler, T.C. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* **2010**, *20*, 45–58. [[CrossRef](#)]
14. Marquez, Y.; Brown, J.W.; Simpson, C.; Barta, A.; Kalyna, M. Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res.* **2012**, *22*, 1184–1195. [[CrossRef](#)]
15. Lu, T.; Lu, G.; Fan, D.; Zhu, C.; Li, W.; Zhao, Q.; Feng, Q.; Zhao, Y.; Guo, Y.; Li, W.; et al. Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res.* **2010**, *20*, 1238–1249. [[CrossRef](#)]
16. Staiger, D.; Brown, J.W. Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* **2013**, *25*, 3640–3656. [[CrossRef](#)]
17. Wu, X.; Liu, M.; Downie, B.; Liang, C.; Ji, G.; Li, Q.Q.; Hunt, A.G. Genome-wide landscape of polyadenylation in *Arabidopsis* provides evidence for extensive alternative polyadenylation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12533–12538. [[CrossRef](#)]
18. Shen, Y.; Ji, G.; Haas, B.J.; Wu, X.; Zheng, J.; Reese, G.J.; Li, Q.Q. Genome level analysis of rice mRNA 3'-end processing signals and alternative polyadenylation. *Nucleic Acids Res.* **2008**, *36*, 3150–3161. [[CrossRef](#)]
19. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* **2015**, *13*, 278–289. [[CrossRef](#)]
20. Xu, Z.; Peters, R.J.; Weirather, J.; Luo, H.; Liao, B.; Zhang, X.; Zhu, Y.; Ji, A.; Zhang, B.; Hu, S.; et al. Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of *Salvia miltiorrhiza* and tanshinone biosynthesis. *Plant J.* **2015**, *82*, 951–961. [[CrossRef](#)]
21. Wang, B.; Tseng, E.; Regulski, M.; Clark, T.A.; Hon, T.; Jiao, Y.; Lu, Z.; Olson, A.; Stein, J.C.; Ware, D. Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. *Nat. Commun.* **2016**, *7*, 11708. [[CrossRef](#)] [[PubMed](#)]
22. Abdel-Ghany, S.E.; Hamilton, M.; Jacobi, J.L.; Ngam, P.; Devitt, N.; Schilkey, F.; Ben-Hur, A.; Reddy, A.S.N. A survey of the sorghum transcriptome using single-molecule long reads. *Nat. Commun.* **2016**, *7*, 11706. [[CrossRef](#)] [[PubMed](#)]
23. Li, Y.; Dai, C.; Hu, C.; Liu, Z.; Kang, C. Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry. *Plant J.* **2017**, *90*, 164–176. [[CrossRef](#)] [[PubMed](#)]
24. Wang, T.; Wang, H.; Cai, D.; Gao, Y.; Zhang, H.; Wang, Y.; Lin, C.; Ma, L.; Gu, L. Comprehensive profiling of rhizome-associated alternative splicing and alternative polyadenylation in moso bamboo (*Phyllostachys edulis*). *Plant J.* **2017**, *91*, 684–699. [[CrossRef](#)] [[PubMed](#)]
25. Chao, Y.; Yuan, J.; Li, S.; Jia, S.; Han, L.; Xu, L. Analysis of transcripts and splice isoforms in red clover (*Trifolium pratense* L.) by single-molecule long-read sequencing. *BMC Plant Biol.* **2018**, *18*, 300. [[CrossRef](#)]
26. Wang, L.; Jiang, X.; Wang, L.; Wang, W.; Fu, C.; Yan, X.; Geng, X. A survey of transcriptome complexity using PacBio single-molecule real-time analysis combined with Illumina RNA sequencing for a better understanding of ricinoleic acid biosynthesis in *Ricinus communis*. *BMC Genom.* **2019**, *20*, 456. [[CrossRef](#)]
27. Ye, J.; Cheng, S.; Zhou, X.; Chen, Z.; Kim, S.U.; Tan, J.; Zheng, J.; Xu, F.; Zhang, W.; Liao, Y.; et al. A global survey of full-length transcriptome of *Ginkgo biloba* reveals transcript variants involved in flavonoid biosynthesis. *Ind. Crop. Prod.* **2019**, *139*, 111547. [[CrossRef](#)]
28. Chao, Q.; Gao, Z.F.; Zhang, D.; Zhao, B.G.; Dong, F.Q.; Fu, C.; Liu, L.; Wang, B. The developmental dynamics of the *Populus* stem transcriptome. *Plant Biotechnol. J.* **2019**, *17*, 206–219. [[CrossRef](#)]
29. Deng, H.; Zhang, L.; Liao, M.; Wang, J.; Liang, D.; Xia, H.; Lv, X.; Deng, Q.; Wang, X.; Tang, Y.; et al. A PacBio single molecule real-time sequencing-based full-length transcriptome atlas of tree tomato (*Solanum betaceum* Cav.) and mining of simple sequence repeat markers. *Front. Plant Sci.* **2022**, *13*, 1052817. [[CrossRef](#)]
30. Yu, Y.; Guo, S.; Ren, Y.; Zhang, J.; Li, M.; Tian, S.; Wang, J.; Sun, H.; Zuo, Y.; Chen, Y.; et al. Quantitative Transcriptomic and Proteomic Analysis of Fruit Development and Ripening in Watermelon (*Citrullus lanatus*). *Front. Plant Sci.* **2022**, *13*, 818392. [[CrossRef](#)]

31. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)] [[PubMed](#)]
32. Campbell, M.A.; Haas, B.J.; Hamilton, J.P.; Mount, S.M.; Buell, C.R. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genom.* **2006**, *7*, 327. [[CrossRef](#)] [[PubMed](#)]
33. Reddy, A.S. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.* **2007**, *58*, 267–294. [[CrossRef](#)] [[PubMed](#)]
34. Yang, J.; Lv, W.; Shao, L.; Fu, Y.; Liu, H.; Yang, C.; Chen, A.; Xie, X.; Wang, Z.; Li, C. PacBio and Illumina RNA Sequencing Identify Alternative Splicing Events in Response to Cold Stress in Two Poplar Species. *Front. Plant Sci.* **2021**, *12*, 737004. [[CrossRef](#)]
35. Xu, W.; Yang, T.; Wang, B.; Han, B.; Zhou, H.; Wang, Y.; Li, D.-Z.; Liu, A. Differential expression networks and inheritance patterns of long non-coding RNAs in castor bean seeds. *Plant J.* **2018**, *95*, 324–340. [[CrossRef](#)]
36. Sun, Y.; Xiao, H. Identification of alternative splicing events by RNA sequencing in early growth tomato fruits. *BMC Genom.* **2015**, *16*, 948. [[CrossRef](#)]
37. Tang, W.; Zheng, Y.; Dong, J.; Yu, J.; Yue, J.; Liu, F.; Guo, X.; Huang, S.; Wisniewski, M.; Sun, J.; et al. Comprehensive Transcriptome Profiling Reveals Long Noncoding RNA Expression and Alternative Splicing Regulation during Fruit Development and Ripening in Kiwifruit (*Actinidia chinensis*). *Front. Plant Sci.* **2016**, *7*, 335. [[CrossRef](#)]
38. Guodong, R.; Jianguo, Z.; Xiaoxia, L.; Ying, L. Identification of putative genes for polyphenol biosynthesis in olive fruits and leaves using full-length transcriptome sequencing. *Food Chem.* **2019**, *300*, 125246. [[CrossRef](#)]
39. Yan, X.; Bai, D.; Song, H.; Lin, K.; Pang, E. Alternative splicing during fruit development among fleshy fruits. *BMC Genom.* **2021**, *22*, 762. [[CrossRef](#)]
40. Gupta, V.; Estrada, A.D.; Blakley, I.; Reid, R.; Patel, K.; Meyer, M.D.; Andersen, S.U.; Brown, A.F.; Lila, M.A.; Loraine, A.E. RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. *Gigascience* **2015**, *4*, 5. [[CrossRef](#)]
41. Terol, J.; Tadeo, F.; Ventimilla, D.; Talon, M. An RNA-Seq-based reference transcriptome for Citrus. *Plant Biotechnol. J.* **2016**, *14*, 938–950. [[CrossRef](#)] [[PubMed](#)]
42. Jiang, J.; Liu, X.; Liu, C.; Liu, G.; Li, S.; Wang, L. Integrating Omics and Alternative Splicing Reveals Insights into Grape Response to High Temperature. *Plant Physiol.* **2017**, *173*, 1502–1518. [[CrossRef](#)]
43. Xanthopoulou, A.; Moysiadis, T.; Bazakos, C.; Karagiannis, E.; Karamichali, I.; Stamatakis, G.; Samiotaki, M.; Manioudaki, M.; Michailidis, M.; Madesis, P.; et al. The perennial fruit tree proteogenomics atlas: A spatial map of the sweet cherry proteome and transcriptome. *Plant J.* **2022**, *109*, 1319–1336. [[CrossRef](#)]
44. Yuan, H.; Yu, H.; Huang, T.; Shen, X.; Xia, J.; Pang, F.; Wang, J.; Zhao, M. The complexity of the *Fragaria x ananassa* (octoploid) transcriptome by single-molecule long-read sequencing. *Hortic. Res.* **2019**, *6*, 46. [[CrossRef](#)] [[PubMed](#)]
45. Saminathan, T.; Nimmakayala, P.; Manohar, S.; Malkaram, S.; Almeida, A.; Cantrell, R.; Tomason, Y.; Abburi, L.; Rahman, M.A.; Vajja, V.G.; et al. Differential gene expression and alternative splicing between diploid and tetraploid watermelon. *J. Exp. Bot.* **2015**, *66*, 1369–1385. [[CrossRef](#)] [[PubMed](#)]
46. Nagalakshmi, U.; Wang, Z.; Waern, K.; Shou, C.; Raha, D.; Gerstein, M.; Snyder, M. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **2008**, *320*, 1344–1349. [[CrossRef](#)] [[PubMed](#)]
47. Fu, H.; Yang, D.; Su, W.; Ma, L.; Shen, Y.; Ji, G.; Ye, X.; Wu, X.; Li, Q.Q. Genome-wide dynamics of alternative polyadenylation in rice. *Genome Res.* **2016**, *26*, 1753–1760. [[CrossRef](#)]
48. Sherstnev, A.; Duc, C.; Cole, C.; Zacharaki, V.; Hornyik, C.; Ozsolak, F.; Milos, P.M.; Barton, G.J.; Simpson, G.G. Direct sequencing of Arabidopsis thaliana RNA reveals patterns of cleavage and polyadenylation. *Nat. Struct. Mol. Biol.* **2012**, *19*, 845–852. [[CrossRef](#)] [[PubMed](#)]
49. Lei, M.; La, H.; Lu, K.; Wang, P.; Miki, D.; Ren, Z.; Duan, C.-G.; Wang, X.; Tang, K.; Zeng, L.; et al. Arabidopsis EDM2 promotes IBM1 distal polyadenylation and regulates genome DNA methylation patterns. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 527–532. [[CrossRef](#)]
50. Wang, X.; Duan, C.G.; Tang, K.; Wang, B.; Zhang, H.; Lei, M.; Lu, K.; Mangrauthia, S.K.; Wang, P.; Zhu, G.; et al. RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intronic heterochromatin-containing gene IBM1. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15467–15472. [[CrossRef](#)]
51. Au, K.F.; Underwood, J.G.; Lee, L.; Wong, W.H. Improving PacBio long read accuracy by short read alignment. *PLoS ONE* **2012**, *7*, e46679. [[CrossRef](#)]
52. Kovaka, S.; Zimin, A.V.; Pertea, G.M.; Razaghi, R.; Salzberg, S.L.; Pertea, M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **2019**, *20*, 278. [[CrossRef](#)]
53. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [[CrossRef](#)]

54. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omic* **2012**, *16*, 284–287. [[CrossRef](#)]
55. Trincado, J.L.; Entizne, J.C.; Hysenaj, G.; Singh, B.; Skalic, M.; Elliott, D.J.; Eyras, E. SUPPA2: Fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol.* **2018**, *19*, 40. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.