

Article

Differences in Alternative Splicing between Yellow and Black-Seeded Rapeseed

Ai Lin ^{1,2,†}, Jinqi Ma ^{1,2,†} , Fei Xu ^{1,2}, Wen Xu ^{1,2}, Huanhuan Jiang ^{1,2}, Haoran Zhang ^{1,2}, Cunmin Qu ^{1,2} , Lijuan Wei ^{1,2,*} and Jiana Li ^{1,2,*}

¹ Chongqing Rapeseed Engineering Research Center, College of Agronomy and Biotechnology, Southwest University, Chongqing 400715, China; linaiiii@email.swu.edu.cn (A.L.); mj2014@email.swu.edu.cn (J.M.); swuxf9310955@email.swu.edu.cn (F.X.); xw1243524211@email.swu.edu.cn (W.X.); jh8469259@email.swu.edu.cn (H.J.); zhr840905528@icloud.com (H.Z.); drquncunmin@swu.edu.cn (C.Q.);

² Academy of Agricultural Sciences, Southwest University, Chongqing 400715, China

* Correspondence: ljn1950@swu.edu.cn (L.W.); lijuanwei@swu.edu.cn (J.L.)

† These authors contributed equally to this work.

Received: 2 July 2020; Accepted: 28 July 2020; Published: 31 July 2020



Abstract: Yellow seed coat color is a desirable characteristic in rapeseed (*Brassica napus*), as it is associated with higher oil content and higher quality of meal. Alternative splicing (AS) is a vital post-transcriptional regulatory process contributing to plant cell differentiation and organ development. To identify novel transcripts and differences at the isoform level that are associated with seed color in *B. napus*, we compared 31 RNA-seq libraries of yellow- and black-seeded *B. napus* at five different developmental stages. AS events in the different samples were highly similar, and intron retention accounted for a large proportion of the observed AS pattern. AS mainly occurred in the early and middle stage of seed development. Weighted gene co-expression network analysis (WGCNA) identified 23 co-expression modules composed of differentially spliced genes, and we picked out two of the modules whose functions were highly associated with seed color. In the two modules, we found candidate DAS (differentially alternative splicing) genes related to the flavonoid pathway, such as *TT8* (*BnaC09g24870D*), *TT5* (*BnaA09g34840D* and *BnaC08g26020D*), *TT12* (*BnaC06g17050D* and *BnaA07g18120D*), *AHA10* (*BnaA08g23220D* and *BnaC08g17280D*), *CHI* (*BnaC09g50050D*), *BAN* (*BnaA03g60670D*) and *DFR* (*BnaC09g17150D*). Gene *BnaC03g23650D*, encoding RNA-binding family protein, was also identified. The splicing of the candidate genes identified in this study might be used to develop stable, yellow-seeded *B. napus*. This study provides insight into the formation of seed coat color in *B. napus*.

Keywords: alternative splicing; RNA-seq; WGCNA; yellow seed; *Brassica napus*

1. Introduction

Alternative splicing, which results in a single gene coding for a variety of proteins [1], is a crucial mechanism for regulating gene expression. The phenomenon was first observed by Gilbert in 1977 [2], and first characterized in 1981 [3]. During the processing of a gene's primary transcript (pre-mRNA), particular segments of the gene may be included or excluded from the final messenger RNA (mRNA), resulting in alternative transcripts for the gene. Alternative splicing occurs frequently in many organisms, where it greatly increases the diversity of encoded proteins in the genome [4]. For example, the proportion of intron-containing genes undergoing alternative splicing (AS) is over 92% in human (*Homo sapiens*) [5], 60% in *Arabidopsis thaliana* [6], 52% in soybean (*Glycine max*) [7], 40% in maize (*Zea mays*) [8] and 40% in *Gossypium raimondii* [9]. Five basic modes of AS events are generally recognized: exon skipping (ES), mutually exclusive exons (ME), alternative donor site (AD), alternative

acceptor site (AA) and intron retention (IR) [10], although other modes of alternative splicing have been observed. Exon skipping is the most common mechanism of AS in mammals, whereby a particular exon may be included or excluded from a gene's mRNA under some conditions, or in particular tissues [11]. In plants, however, the most frequent AS event is intron retention [7].

mRNA is spliced by the spliceosome, an RNA–protein complex consisting of small nuclear ribonucleoprotein (snRNP) modules and a variety of auxiliary proteins [12]. Splice sites are determined by cis-acting elements, trans-acting RNA-binding proteins (RBPs) and splicing repressor proteins [13] such as exonic splicing enhancers, intronic splicing silencers and so on [14]. A typical eukaryotic nuclear intron contains consensus sequences that are of vital importance for splicing. Nearly all plant introns have the sequence GU at the 5' end. Near the 3' end there is a branch point which is followed by a series of pyrimidines (the poly-pyrimidine tract) and then by AG at the 3' end. This consensus sequence is known as a "GU-AG boundary", and introns containing this sequence follow the "GU-AG" rule [15]. This kind of intron is classified as a U2 type intron, the most common type of intron in eukaryotes (the other being the U12 type) [16].

Alternative splicing of mRNA precursors greatly increases the diversity of mRNA and protein isoforms, and the process influences all aspects of organisms. In plants, alternative splicing contributes to cell differentiation [17], organ development [8], acquisition and maintenance of tissue identity [1] and responses to biotic and abiotic stresses such as heat and disease [18]. It was reported that constitutive and inducible alternative splicing of OsWRKY62 and OsWRKY76 in rice (*Oryza sativa*) may help defend against the blast fungus *Magnaporthe oryzae* and the leaf blight bacterium *Xanthomonas oryzae pv oryzae* [19].

Brassica napus is one of the most important oil crops in the world [20]. Seed coat color is an important trait in *B. napus*. Yellow-seeded *B. napus* is more valuable than black-seeded *B. napus*, because of its higher oil content and better-quality meal [21]. Yellow-seeded *B. napus* also has other advantages over black-seeded varieties: the seed coat is thinner and the oil is clearer, which not only retains the original flavor of the vegetable oil but also gives it a good commercial appearance. To identify the candidate genes for seed coat color, previous researchers mainly performed QTL mapping analysis [22], and focused on the flavonoid biosynthesis pathway [23]. The mutation of genes in the flavonoid synthesis pathway results in defective pigment and proanthocyanidins synthesis, which might form an abnormal or transparent seed coat. Thus, yellow seeds show the color of the embryo due to the colorless seed coat [23]. However, to our knowledge, there have been no post-transcriptional regulatory studies about the formation mechanism of yellow-seeded *B. napus*.

In this study, we analyzed alternative splicing in yellow- and black-seeded *B. napus*, and identified the DAS genes associated with seed color in *B. napus*. The candidate genes involved the flavonoid pathway, and some genes encoded the SER/ARG-rich (SR) protein and RNA-binding family protein.

2. Results

2.1. Putative Transcript Assembly

To understand the role of alternative splicing (AS) in black-seeded and yellow-seeded *B. napus*, 31 RNA-seq data of seeds in five developmental stages after flowering (14, 21, 28, 35 and 42 d) were obtained [24] and subjected to quality trimming, followed by transcript assembly and transcript merging using the RNA-seq data-analysis pipeline. We identified 4249, 5164, 5953, 6341 and 6438 new transcripts in yellow-seeded *B. napus* for the five developmental stages, and 4684, 5747, 6427, 6626 and 7000 in black-seeded *B. napus* with the class code "u", none of which exist in the reference genome (Table S1). We also obtained the FPKM values for the genes and isoforms. The average FPKM values of the yellow- and black-seeded *B. napus* at the five developmental stages were 18.09, 54.60, 20.19, 16.17, 20.45, 12.67, 37.74, 18.10, 73.10 and 250.19, respectively (Table S1), whereas the average FPKM values for the new transcripts were 33.64, 593.20, 85.92, 32.46, 22.04, 16.78, 282.08, 44.89, 564.86 and 645.39. Thus, the gene expression levels were higher for the new transcripts.

2.2. Analysis of the AS Landscape in Yellow- and Black-Seeded *B. napus*

The AS patterns and landscapes of yellow- and black-seeded *B. napus* were identified using Astaravista-4.0. We detected 21,931, 20,142, 21,237, 18,842 and 19,801 AS events in five developmental stages of yellow-seeded *B. napus*, respectively. We also detected 20,248, 20,366, 23,519, 19,806 and 20,471 AS events in five developmental stages of black-seeded *B. napus*, respectively. By comparing the AS patterns, we found that IR is the most common one, followed by AA, AD and ES (Figure 1A, Table S2).

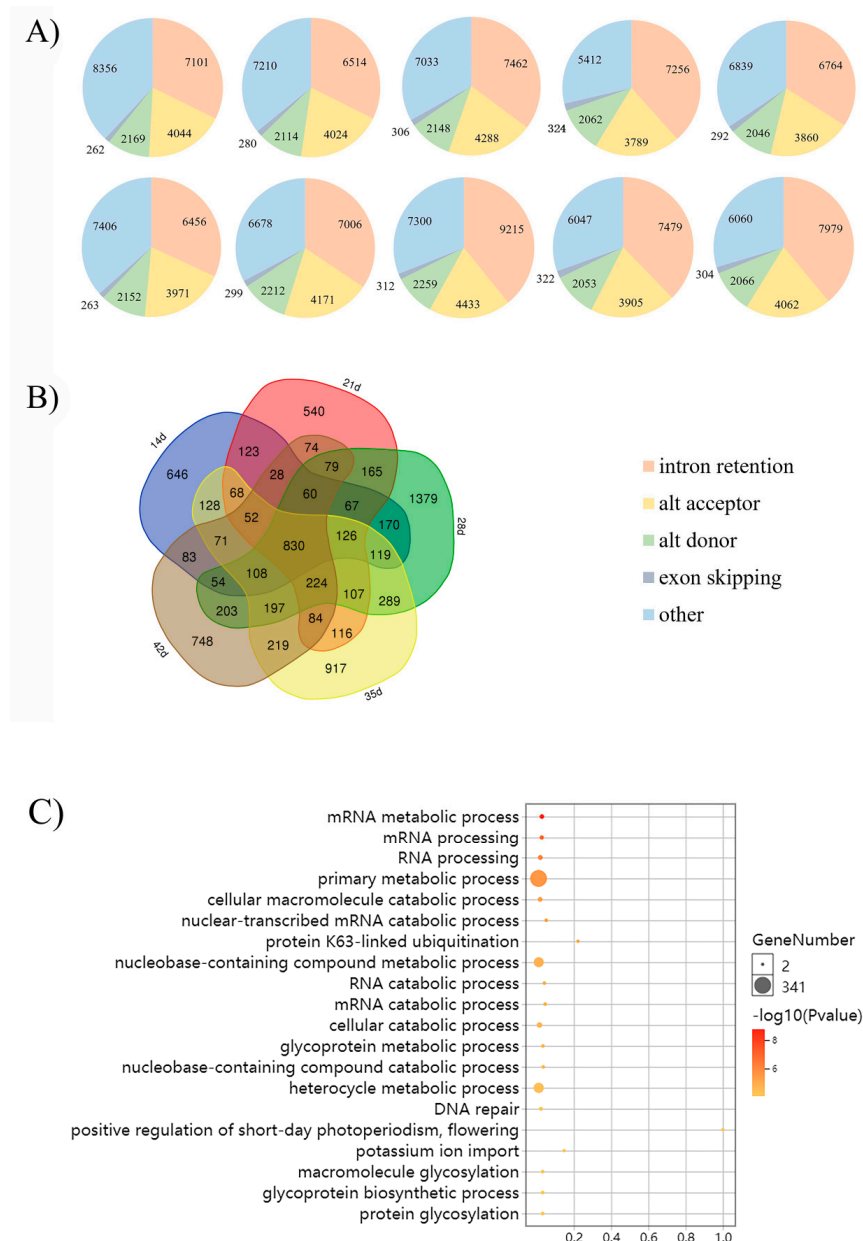


Figure 1. The alternative splicing (AS) landscape of yellow- and black-seeded *B. napus* in five developmental stages of seed development (14, 21, 28, 35 and 42 d after flowering). (A) The AS patterns of genes in yellow- and black-seeded *B. napus* in five developmental stages of seed development (14, 21, 28, 35 and 42 d after flowering). The first row represents data from black-seeded samples, and the second from yellow-seeded samples. (B) Genes exhibiting AS at five developmental stages (14, 21, 28, 35 and 42 d after flowering). The pentagon at the center represents the 830 overlapping differentially spliced genes at the five specific time points. (C) Visualization of 830 DAS (differentially alternative splicing) genes identified using the enriched Gene Ontology (GO) biological process.

The results indicate that the AS events in five developmental stages of yellow- and black-seeded *B. napus* were almost same over the course of seed development, and that IR accounted for a large proportion of the AS events. The AS patterns were highly conserved in all samples.

2.3. Differential AS gene

We detected 2733, 2743, 4177, 3655 and 3144 DAS genes in five developmental stages, 14, 21, 28, 35 and 42 d after flowering, respectively (Figure 1B). A total of 830 overlapping DAS genes in all five developmental stages between yellow- and black-seeded samples were identified (Figure 1B, Table S3). These genes mainly participate in metabolic/catabolic processes, protein ubiquitination, potassium ion import, DNA repair and short-day photoperiodism (Figure 1C). Three DAS genes encoding SER/ARG-rich (SR) protein (*BnaC08g21130D*, *BnaC01g26160D* and *BnaA06g15930D*) were identified, which can bind to some particular RNA sequences and assemble the spliceosome at weak splice sites, and thus play a significant role in the constitutive and alternative splicing of pre-mRNA in rice [25]. We also identified DNA/RNA-binding protein (*BnaC04g41260D*, *BnaAnng15850D*, *BnaC03g02460D* and *BnaC07g01360D*), which has been reported to be involved in circadian clock regulation [26] with an hnRNP-like RNA-binding protein in *A. thaliana* [27]. We also found some transcription factors, such as the basic region/leucine zipper transcription factor 16 (bZIP16).

We also identified 646, 540, 1379, 917 and 748 specific DAS genes in seeds at 14, 21, 28, 35 and 42 d after flowering, respectively. After performing Gene Ontology (GO) analysis of specific DAS genes in five developmental stages of yellow- and black-seeded *B. napus* (Table S4), we observed that in the early and middle stages of seed development (14, 21 and 28 d), the DAS genes were mainly enriched in mRNA splicing. The specific DAS genes in seeds at 35 d after flowering were mainly involved in monocarboxylic acid metabolic process. While at 42 d after flowering, the specific DAS genes were mainly enriched in localization, organophosphate metabolic process, and catabolic processes of aromatic compounds, heterocycles, cellular nitrogen compounds and organic cyclic compounds. The results indicate that AS mainly occurred in the early and middle stages of seed development. In addition, we found DAS genes involved in flavonoid metabolism, such as *transparent testa* genes (TT5: *BnaC02g38340D*, *BnaC08g26020D*, TT8: *BnaC09g24870D*, TT10: *BnaAnng08030D*, *BnaA06g30430D*, *BnaA09g34840D*, TT12: *BnaC06g17050D*, *BnaA07g18120D*, TT16: *BnaAnng30140D* and *BnaC09g04950D*) and *seedstick* (STK) (*BnaAnng39120D* and *BnaCnng46740D*) (Table 1).

Table 1. Candidate DAS genes related to seed coat color in yellow- and black-seeded *B. napus*.

Gene	Homologous Genes	Annotated	Module	Stage
BnaA03g60670D	AT1G61720	BANYULS (BAN)	black	28 d
BnaC09g50050D	AT5G05270	Chalcone-flavanone isomerase family protein (CHI)	black	28 d
BnaC09g17150D	AT5G42800	dihydroflavonol 4-reductase (DFR)	dark grey	28 d, 35 d, 42 d
BnaA08g23220D	AT1G17260	autoinhibited H(+)-ATPase isoform 10 (AHA10)	black	14 d, 21 d, 28 d, 42 d
BnaC08g17280D	AT1G17260	autoinhibited H(+)-ATPase isoform 10 (AHA10)	black	21 d, 28 d
BnaAnng39120D	AT4G09960	SEEDSTICK (STK)	not included	14 d, 21 d, 28 d, 35 d, 42 d
BnaCnng46740D	AT4G09960	SEEDSTICK (STK)	dark red	42 d
BnaC08g21130D	AT3G49430	SER/ARG-rich protein 34A (SRp34a)	not included	14 d, 21 d, 28 d, 35 d, 42 d
BnaC01g26160D	AT3G49430	SER/ARG-rich protein 34A (SRp34a)	not included	14 d, 21 d, 28 d, 35 d, 42 d
BnaA06g15930D	AT3G49430	SER/ARG-rich protein 34A (SRp34a)	turquoise	14 d, 21 d, 28 d, 35 d, 42 d
BnaA09g34840D	AT3G55120	TRANSPARENT TESTA 5 (TT5)	black	28 d, 42 d
BnaC08g26020D	AT3G55120	TRANSPARENT TESTA 5 (TT5)	black	28 d, 35 d, 42 d
BnaC09g24870D	AT4G09820	TRANSPARENT TESTA 8 (TT8)	black	21 d, 28 d, 35 d, 42 d
BnaC06g17050D	AT3G59030	TRANSPARENT TESTA 12 (TT12)	black	21 d, 28 d, 35 d, 42 d
BnaA07g18120D	AT3G59030	TRANSPARENT TESTA 12 (TT12)	black	42 d
BnaAnng08030D	AT5G48100	TRANSPARENT TESTA 10 (TT10)	not included	21 d, 28 d, 35 d, 42 d
BnaA06g30430D	AT5G48100	TRANSPARENT TESTA 10 (TT10)	blue	21 d, 28 d, 35 d, 42 d
BnaC02g38340D	AT5G48100	TRANSPARENT TESTA 10 (TT10)	turquoise	28 d, 35 d, 42 d
BnaAnng30140D	AT5G23260	TRANSPARENT TESTA16 (TT16)	turquoise	14 d
BnaC09g04950D	AT5G23260	TRANSPARENT TESTA16 (TT16)	turquoise	42 d
BnaC04g41260D	AT5G03495	RNA-binding family protein	not included	14 d, 21 d, 28 d, 35 d, 42 d
BnaC07g01360D	AT5G03495	RNA-binding family protein	turquoise	14 d, 21 d, 28 d, 35 d, 42 d
BnaC03g02460D	AT5G03480	RNA-binding family protein	turquoise	14 d, 21 d, 28 d, 35 d, 42 d
BnaAnng15850D	AT5G03480	RNA-binding family protein	blue	14 d, 21 d, 28 d, 35 d, 42 d

2.4. Identification of Modules Associated with Seed Color by Construction of Weighted Gene Co-Expression Network and Correlation between Each Module

To uncover the major candidate genes for seed color, we performed weighted gene co-expression network analysis (WGCNA) with all DAS genes. WGCNA can group similarly expressed genes into modules or networks based on pairwise correlations between genes [28]. Using a soft threshold of 5, the co-expression network was generated and 23 modules were obtained (Figure 2A,B).

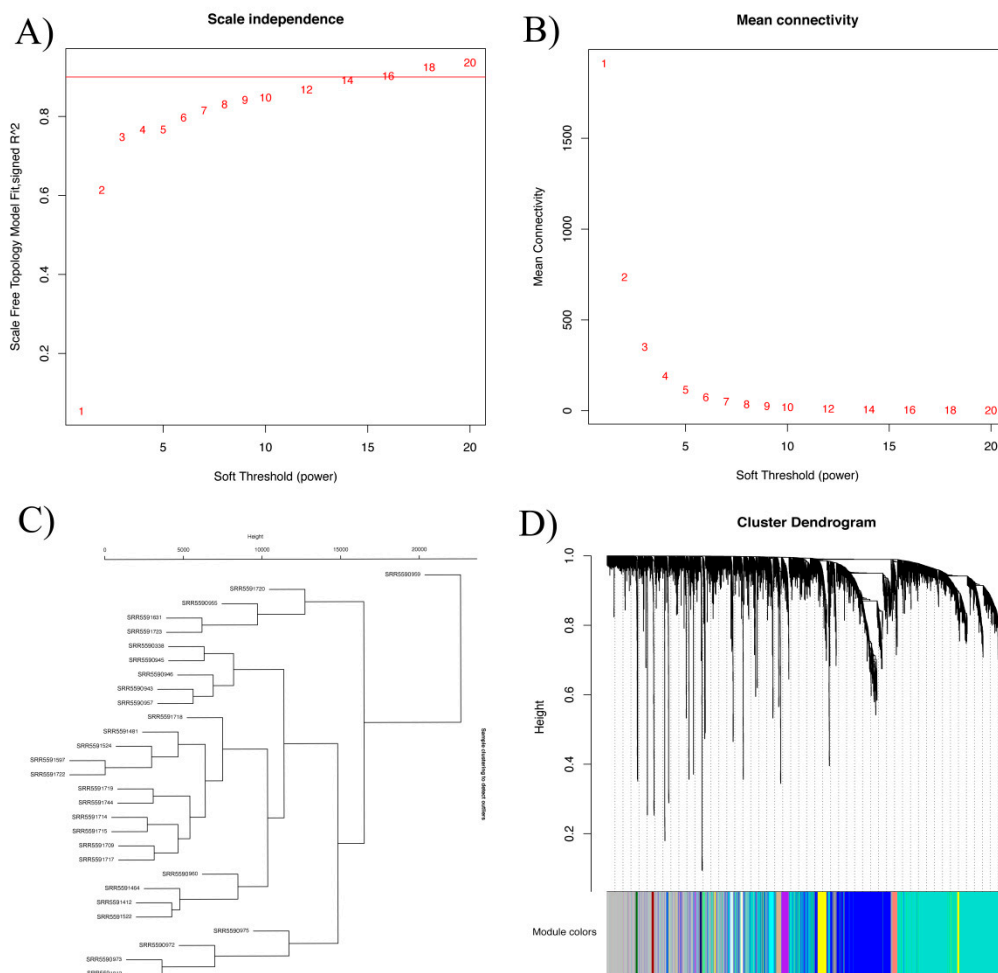


Figure 2. Module identification by weighted gene co-expression network analysis (WGCNA). (A,B) Network topology of the soft threshold with scale independence and mean connectivity. The approximate scale-free topology can be acquired at a soft threshold of 5. (C) Dendrogram of consensus module eigengenes obtained by WGCNA on the consensus correlation. (D) Gene dendrogram obtained by clustering the dissimilarity, based on consensus topological overlap with the corresponding module colors, indicated by the color row. Each colored row represents a color-coded module that contains a group of highly connected genes. A total of 23 modules were identified. The genes in the gray module are not significantly related, and the genes were not classified into any module.

Meanwhile, we performed GO analysis to explore module function, with two modules (black and dark grey modules) related to flavonoid metabolic pathways that have been reported to be important for seed color (Figure S1). The function of the other 21 modules is also shown in Table S5. The genes in the black module are mainly involved in the flavonoid biosynthetic and metabolic process, the proanthocyanidin metabolic process, the fatty acid catabolic process and the maintenance of seed dormancy. The genes in the dark grey module are mainly involved in the anthocyanin-containing

compound biosynthetic process, the fatty-acyl-CoA metabolic process, and the fatty acid derivative metabolic process.

Using a weight value of >0.15 , we performed a Cytoscape network analysis for the two vital modules to show the relationships among genes in a single module (Figure 3 and Table S6). The black and dark grey modules contained 171 and 25 genes, respectively, and the most and least highly connected genes are shown in red and purple, respectively.

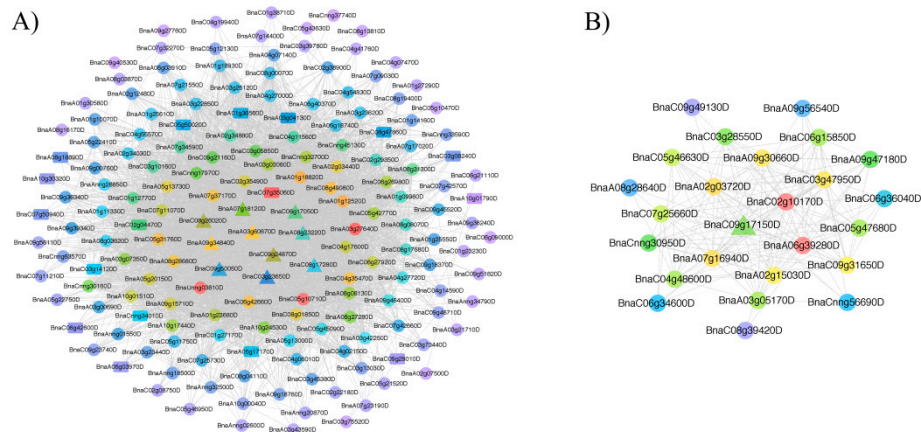


Figure 3. Visualization of two vital co-expression networks identified using Cytoscape. (A) the black module and (B) dark grey module. The candidate DAS genes are indicated by triangles, the intersection of genes in two vital functional modules and 830 overlapping genes are indicated by rectangles, whereas the remaining genes are indicated by circles.

Among the black and dark grey modules, the candidate DAS genes involved in flavonoid metabolism were identified, such as TT8 (*transparent testa 8*, *BnaC09g24870D*), TT5 (*BnaA09g34840D* and *BnaC08g2620D*), TT12 (*BnaC06g17050D* and *BnaA07g18120D*), AHA10 (autoinhibited H(+)-ATPase isoform 10, *BnaA08g23220D* and *BnaC08g17280D*), CHI (chalcone-flavanone isomerase family protein, *BnaC09g50050D*), BAN (BANYULS, *BnaA03g60670D*) and DFR (dihydroflavonol 4-reductase, *BnaC09g17150D*) (Figure 3 and Table 1). In addition, we found one gene (*BnaC03g23650D*) that encodes RNA-binding family protein. These DAS genes might play a major role in the formation of seed coat color.

3. Discussion

In this study, we revealed the AS landscape and identified DAS genes between yellow-seeded and black-seeded oilseed rape. We performed GO analysis on 830 overlapping markedly DAS genes. The major GO terms of these genes include transport, regulation of protein homodimerization activity, establishment of localization, organic substance transport and exocyst assembly.

In order to identify the candidate genes involved in seed color, we used WGCNA to classify the DAS genes into 23 modules. Among the 23 modules, we performed GO analyses and module-sample cluster analysis to choose two modules related with seed coat color formation. It has been reported that pigments and other chemical substances affect the seed coat color during seed coat formation [23]. Flavonoids are a class of low-molecular-weight phenolic compounds, and are representative plant secondary products. In addition, flavonoids play a vital role in sexual reproduction and in protecting plants from UV damage [29]. The chemical diversity of flavonoids can be increased enormously by tailoring reactions that modify scaffolds, including glycosylation, methylation and acylation [30]. Based on the above pathways, we chose modules involved in flavonoid, lignin, phenylpropanoid biosynthetic and metabolic processes, as well as pigmentation and anthocyanin accumulation, all of which correspond to previously reported results [23,31].

Alternative splicing in flavonoid genes or transcription factors (TFs) affect the color of the plant. In peach flowers, the *ANS* gene of white flowers performed intron retention of a spare 129 bp sequence lacking in pink flowers, which leads to the dysfunction of the protein, and finally blocks the accumulation of colored pigments in white petals [32]. In tomato, splicing mutations of transcription-factor-encoding genes which control the loci of anthocyanin pigmentation leads to a loss of function of the wild-type protein. This is why domesticated tomato fruits display a red coloration instead of the purple color in wild species [33]. According to the functional annotation of *Arabidopsis*, homologous genes of 11 DAS genes were related to the seed color: TT8, two TT5 genes, two TT12, two AHA10, CHI, BAN, DFR and *BnaC03g23650D*. TT8 encodes an enzyme that regulates the flavonoid pathway [34] and a TF involved in the regulation of anthocyanin [35]. The expression of TT8 is strongly connected with seed coat differentiation in *Arabidopsis thaliana* [34]. TT5 is involved in the synthesis of chalcones in the flavonoid pathway [36]. TT12, a kind of transporter of proanthocyanidins (PAs) [37], is suggested to regulate the accumulation of proanthocyanidin in the seed [38]. It has been shown that the relative expression of TT12 in brown cotton was higher than in white cotton [37]. AHA10 has been proven to be the early biosynthetic gene involved in flavonoid biosynthesis [23]. Highly expressed AHA10 was involved in the acidification of lemon (*Citrus limon* (L.) Burm.) [39]. CHI encodes the second enzyme of flavonoid biosynthesis; it catalyzes chalcone into corresponding flavones [40]. Research has shown that BAN is closely associated with *Arabidopsis thaliana* seed coat [41]; it codes the central enzyme of proanthocyanidin biosynthesis [35]. DFR is the late biosynthetic gene of flavonoid biosynthesis; it participates in catalytic synthesis of anthocyanin and procyanidins in plants [42]. The DAS genes found in this study might play a major role in the formation of seed coat color in *B. napus*, which needs to be verified in the future.

We also selected some crucial AS genes, such as *STK*, genes encoding SR protein and TF bZIP6. *STK*, a transcription factor that could facilitate initiation and growth of ovule [43], has also been proven to regulate cell wall properties of the seed coat [44]. SR protein plays vital roles in the constitutive and alternative splicing of pre-mRNA [25]. It has been proven to enhance splicing and alter the splice sites in rice, showing the significance of the SR domain for the enhancement of splicing efficiencies [25], thus regulating the growth and development of plants. Research has shown that bZIP16 promotes seed germination and hypocotyl elongation by repressing the expression of some responsive hormone genes [45]. In this study, we explored candidate genes related to seed color through AS, thereby providing more possibilities to develop stable yellow-seeded *B. napus*.

4. Materials and Methods

4.1. Downloading of RNA-Seq Data

We downloaded RNA-seq data from the seeds of yellow- (cultivar: No.2127-17) and black-seeded (cultivar: 94,570) *B. napus* in five developmental stages, 14, 21, 28, 35 and 42 d seeds after flowering, with three or four biological replicates: SRR5590338, SRR5590943, SRR5590945 and SRR5591364 (14 d; black-seeded); SRR5590946, SRR5590955 and SRR5590957 (14 d; yellow-seeded); SRR5590959, SRR5590960 and SRR5590972 (21 d; black-seeded); SRR5590973, SRR5590975 and SRR5591012 (21 d; yellow-seeded); SRR5591412, SRR5591464 and SRR5591481 (28 d; black-seeded); SRR5591522, SRR5591524 and SRR5591597 (28 d; yellow-seeded); SRR5591631, SRR5591634 and SRR5591709 (35 d; black-seeded); SRR5591714, SRR5591715 and SRR5591717 (35 d; yellow-seeded); SRR5591718, SRR5591719 and SRR5591720 (42 d; black-seeded); and SRR5591722, SRR5591723 and SRR5591744 (42 d; yellow-seeded) [46].

4.2. Analysis of Processed RNA-Seq Data and AS Landscape

The original data were converted using the SRA Toolkit 2.9.0 [47], and quality trimming was performed with Trimmomatic-0.36 [48] (modified parameters: ILLUMINACLIP: TruSeq3-PE.fa:2:30:10, LEADING:3, TRAILING:3, and SLIDINGWINDOW:4:15, MINLEN:50). Then, the processed groups of

data were mapped to the *B. napus* reference genome [49] using STAR-2.5.3a [50] (with `sjdbOverhang` = 150 and `limitBAMsortRAM` = 5,189,162,595), and Cufflinks-2.2.1 was used to assemble putative transcripts [51]. The subsequent program Cuffcompare could map the different isoforms to the corresponding genes [52], and the last step was to analyze the AS patterns and to visualize the AS landscape using Astalavista-4.0 [53].

4.3. Construction of Weighted Gene Co-Expression Networks and Analysis of Overlapping Differentially Spliced Genes

WGCNA, a freely accessible R package for the construction of weighted gene co-expression networks [54], can help identify gene modules of interest from thousands of seed coat genes. WGCNA in R version 3.4.4 was used to perform sample clustering, outlier detection, soft threshold filtering, module identification and module relationship analysis. In our study, the one-step function was used to construct networks and detect consensus modules.

4.4. Analysis of Overlapping Differentially Spliced Genes

We used rMATS [55] under a Python console to investigate the differential alternative splicing between yellow- and black-seeded samples. The rMATS method allows analysis of all major types of alternative splicing patterns and uses RNA-Seq reads, mapped to both exons and splice junctions [55]. Fastq format files were uploaded and the STAR index file of the genome was used for the comparative analysis.

4.5. Functional Enrichment Analysis of Genes in Every Module

The genes inside the co-expression modules identified by WGCNA have high connectivity and similar function. We performed the R package topGO to perform the GO analysis [56]. We filtered the hub genes in each module according to the intra-modular connectivity and further understood their function.

5. Conclusions

In this study, we identified candidate alternative splicing genes between yellow-seeded and black-seeded oilseed rape during the post-transcriptional regulatory process. The function of the candidate genes needs to be verified by molecular biology experiments in the future. This study will help us understand the formation mechanism of yellow-seeded *B. napus*.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2223-7747/9/8/977/s1>, Figure S1: The enriched GO biological process categories of black and dark grey modules. Boxes indicate the 10 most significant terms. Box color represents significance, ranging from dark red (most significant) to light yellow (least significant). Black arrows indicate hierarchical relationships of the GO term, and the arrow direction represents the containment relationship. Table S1: FPKM values of all transcripts of yellow- and black-seeded samples on five specific days after flowering, with three replicates. Table S2: AS patterns of various plant materials. Table S3: Enriched GO terms of genes and their functions of the 830 differentially spliced genes and their FPKM values in yellow- and black-seeded samples. Table S4: Enriched GO terms of genes and their functions of specific DAS genes in seeds at 5 stages after flowering. Table S5: Enriched GO terms of genes and their functions of 23 modules. Table S6: Network relationships of black and dark grey modules.

Author Contributions: Conceptualization, J.L.; Methodology, A.L. and J.M.; Writing—Original Draft Preparation, A.L. and H.Z.; Investigation, F.X., W.X.; Writing—Review and Editing, A.L., J.M., L.W., H.J.; Funding Acquisition, J.L., L.W., C.Q., and A.L. All authors have read and agreed to the published version of the manuscript.

Funding: Please add: This research was funded by grants from the National Natural Science Foundation of China (31830067), National Key Research and Development Program of China (2018YFD0100504-05), and the “111” Project (B12006).

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

References

1. Baralle, F.E.; Giudice, J. Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 437. [[CrossRef](#)]
2. Chow, L.T.; Gelinis, R.E.; Broker, T.R.; Roberts, R.J. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **1977**, *12*, 1–8. [[CrossRef](#)]
3. DeNoto, F.M.; Moore, D.D.; Goodman, H.M. Human growth hormone DNA sequence and mRNA structure: Possible alternative splicing. *Nucleic Acids Res.* **1981**, *9*, 3719–3730. [[CrossRef](#)] [[PubMed](#)]
4. Black, D.L. Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* **2003**, *72*, 291–336. [[CrossRef](#)] [[PubMed](#)]
5. Wang, E.T.; Rickard, S.; Shujun, L.; Irina, K.; Lu, Z.; Christine, M.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008**, *456*, 470–476. [[CrossRef](#)] [[PubMed](#)]
6. 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](#)]
7. Shen, Y.; Zhou, Z.; Wang, Z.; Li, W.; Fang, C.; Wu, M.; Ma, Y.; Liu, T.; Kong, L.-A.; Peng, D.-L. Global Dissection of Alternative Splicing in Paleopolyploid Soybean. *Plant Cell* **2014**, *26*, 996. [[CrossRef](#)]
8. Thatcher, S.R.; Zhou, W.; Leonard, A.; Wang, B.B.; Beatty, M.; Zastrow-Hayes, G.; Zhao, X.; Baumgarten, A.; Li, B. Genome-wide analysis of alternative splicing in *Zea mays*: Landscape and genetic regulation. *Plant Cell* **2014**, *26*, 3472–3487. [[CrossRef](#)]
9. Li, Q.; Xiao, G.; Zhu, Y. Single-Nucleotide Resolution Mapping of the *Gossypium raimondii* Transcriptome Reveals a New Mechanism for Alternative Splicing of Introns. *Mol. Plant* **2014**, *7*, 829. [[CrossRef](#)]
10. Pan, Q.; Shai, O.; Lee, L.J.; Frey, B.J.; Blencowe, B.J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **2008**, *40*, 1413–1415. [[CrossRef](#)]
11. Sammeth, M.; Foissac, S.; Guigo, R. A general definition and nomenclature for alternative splicing events. *PLoS Comput. Biol.* **2008**, *4*, e1000147. [[CrossRef](#)] [[PubMed](#)]
12. Mo, C.; Manley, J.L. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 540. [[CrossRef](#)]
13. Luo, Z.; Yang, Q.; Yang, L. RNA Structure Switches RBP Binding. *Mol. Cell* **2016**, *64*, 219–220. [[CrossRef](#)] [[PubMed](#)]
14. Xiang-Dong, F.; Manuel, A.J. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **2014**, *15*, 689–701. [[CrossRef](#)]
15. Matlin, A.J.; Francis, C.; Smith, C.W.J. Understanding alternative splicing: Towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 386–398. [[CrossRef](#)]
16. Lazarescu, E.; Wolfgang, F.; Andre, S. Organ-specific alternatively spliced transcript isoforms of the sunflower SF21C gene. *Plant Cell Rep.* **2010**, *29*, 673. [[CrossRef](#)]
17. Fiszbein, A.; Kornblihtt, A.R. Alternative splicing switches: Important players in cell differentiation. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **2017**, *39*, 1600157. [[CrossRef](#)]
18. Chiung-Yun, C.; Wen-Dar, L.; Shih-Long, T. Genome-Wide Analysis of Heat-Sensitive Alternative Splicing in *Physcomitrella patens*. *Plant Physiol.* **2014**, *165*, 826. [[CrossRef](#)]
19. Liu, J.; Chen, X.; Linag, X.; Zhou, X.; Yang, F.; Liu, J.; He, S.; Guo, Z. Alternative Splicing of Rice WRKY62 and WRKY76 Transcription Factor Genes in Pathogen Defense. *Plant Physiol.* **2016**, *171*, 1427–1442. [[CrossRef](#)]
20. Wen, J.; Zhu, L.; Qi, L.; Ke, H.; Yi, B.; Shen, J.; Tu, J.; Ma, C.; Fu, T. Characterization of interplod hybrids from crosses between *Brassica juncea* and *B. oleracea* and the production of yellow-seeded *B. napus*. *Theor. Appl. Genet.* **2012**, *125*, 19–32. [[CrossRef](#)]
21. Akhov, L.; Ashe, P.; Tan, Y.; Datla, R.; Selvaraj, G.; Peterson, R.J.; Shelp, B.J. Proanthocyanidin biosynthesis in the seed coat of yellow-seeded, canola quality *Brassica napus* YN01-429 is constrained at the committed step catalyzed by dihydroflavonol 4-reductase. *Bot.-Bot.* **2009**, *87*, 616–625. [[CrossRef](#)]
22. Badani, A.G.; Snowdon, R.J.; Wittkop, B.; Lipsa, F.D.; Baetzel, R.; Horn, R.; De Haro, A.; Font, R.; Luhs, W.; Friedt, W. Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). *Genome* **2006**, *49*, 1499–1509. [[CrossRef](#)] [[PubMed](#)]

23. Lepiniec, L.; Debeaujon, I.; Routaboul, J.M.; Baudry, A.; Pourcel, L.; Nesi, N.; Caboche, M. Genetics and Biochemistry of Seed Flavonoids. *Annu. Rev. Plant Biol.* **2006**, *57*, 405–430. [[CrossRef](#)] [[PubMed](#)]
24. Jiang, J.; Zhu, S.; Yuan, Y.; Wang, Y.; Zeng, L.; Batley, J.; Wang, Y. Transcriptomic comparison between developing seeds of yellow- and black-seeded Brassica napus reveals that genes influence seed quality. *BMC Plant Biol.* **2019**, *19*, 203. [[CrossRef](#)] [[PubMed](#)]
25. Isshiki, M.; Tsumoto, A.; Shimamoto, K. The serine/arginine-rich protein family in rice plays important roles in constitutive and alternative splicing of pre-mRNA. *Plant Cell* **2006**, *18*, 146–158. [[CrossRef](#)]
26. Staiger, D.; Zecca, L.; Kirk, D.A.W.; Apel, K.; Eckstein, L. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. *Plant J.* **2003**, *33*, 361–371. [[CrossRef](#)]
27. Streitner, C.; Koster, T.; Simpson, C.G.; Shaw, P.; Staiger, D. An hnRNP-like RNA-binding protein affects alternative splicing by in vivo interaction with transcripts in Arabidopsis thaliana. *Nucleic Acids Res.* **2012**, *40*, 11240–11255. [[CrossRef](#)] [[PubMed](#)]
28. Li, Y.; Cai, Z.; Zhu, B.; Xu, C. Identification of Key Pathways and Genes in the Dynamic Progression of HCC Based on WGCNA. *Genes* **2018**, *9*, 92. [[CrossRef](#)]
29. Koes, R.E.; Quattrocchio, F.; Mol, J.N.M. The flavonoid biosynthetic pathway in plants: Function and evolution. *BioEssays* **1994**, *16*, 123–132. [[CrossRef](#)]
30. Saito, K.; Yonekura-Sakakibara, K.; Nakabayashi, R.; Higashi, Y.; Yamazaki, M.; Tohge, T.; Fernie, A.R. The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity. *Plant Physiol. Biochem. PPB* **2013**, *72*, 21–34. [[CrossRef](#)]
31. Wei, Y.-L.; Li, J.-N.; Lu, J.; Tang, Z.-L.; Pu, D.-C.; Chai, Y.-R. Molecular cloning of Brassica napus TRANSPARENT TESTA 2 gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. *Mol. Biol. Rep.* **2007**, *34*, 105–120. [[CrossRef](#)] [[PubMed](#)]
32. Yin, P.; Zhen, Y.; Li, S. Identification and functional classification of differentially expressed proteins and insight into regulatory mechanism about flower color variegation in peach. *Acta Physiol. Plant.* **2019**, *41*, 95. [[CrossRef](#)]
33. Colanero, S.; Tagliani, A.; Perata, P.; Gonzali, S. Alternative Splicing in the Anthocyanin Fruit Gene Encoding an R2R3 MYB Transcription Factor Affects Anthocyanin Biosynthesis in Tomato Fruits. *Plant Commun.* **2020**, *1*, 100006. [[CrossRef](#)]
34. Xu, W.; Grain, D.; Le Gourrierel, J.; Harscoet, E.; Berger, A.; Jauvion, V.; Scagnelli, A.; Berger, N.; Bidzinski, P.; Kelemen, Z. Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of TT8 expression, in Arabidopsis. *New Phytol.* **2013**, *198*, 59–70. [[CrossRef](#)] [[PubMed](#)]
35. Baudry, A.; Heim, M.A.; Dubreucq, B.; Caboche, M.; Weisshaar, B.; Lepiniec, L. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. *Plant J.* **2004**, *39*, 366–380. [[CrossRef](#)] [[PubMed](#)]
36. Böttcher, C.; von Roepenack-Lahaye, E.; Schmidt, J.; Schmotz, C.; Neumann, S.; Scheel, D.; Clemens, S. Metabolome Analysis of Biosynthetic Mutants Reveals a Diversity of Metabolic Changes and Allows Identification of a Large Number of New Compounds in Arabidopsis. *Plant Physiol.* **2008**, *147*, 2107–2120. [[CrossRef](#)]
37. Gao, J.-S.; Wu, N.; Shen, Z.-L.; Lv, K.; Qian, S.-H.; Guo, N.; Sun, X.; Cai, Y.-P.; Lin, Y. Molecular cloning, expression analysis and subcellular localization of a Transparent Testa 12 ortholog in brown cotton (*Gossypium hirsutum* L.). *Gene* **2016**, *576 Pt 2*, 763–769. [[CrossRef](#)]
38. Marinova, K.; Pourcel, L.; Weder, B.; Schwarz, M.; Barron, D.; Routaboul, J.M.; Debeaujon, I.; Klein, M. The Arabidopsis MATE transporter TT12 acts as a vacuolar flavonoid/H⁺ -antiporter active in proanthocyanidin-accumulating cells of the seed coat. *Plant Cell* **2007**, *19*, 2023–2038. [[CrossRef](#)] [[PubMed](#)]
39. Aprile, A.; Federici, C.; Close, T.J.; De Bellis, L.; Cattivelli, L.; Roose, M.L. Expression of the H⁺-ATPase AHA10 proton pump is associated with citric acid accumulation in lemon juice sac cells. *Funct. Integr. Genom.* **2011**, *11*, 551–563. [[CrossRef](#)] [[PubMed](#)]
40. Van Tunen, A.J.; Kose, R.E.; Spelt, C.E.; van der Krol, A.R.; Stuitje, A.R.; Mol, J.N. Cloning of the two chalcone flavanone isomerase genes from Petunia hybrida: Coordinate, light-regulated and differential expression of flavonoid genes. *EMBO J.* **1988**, *7*, 1257–1263. [[CrossRef](#)]
41. Xie, D.-Y.; Sharma, S.B.; Paiva, N.L.; Ferreira, D.; Dixon, R.A. Role of Anthocyanidin Reductase, Encoded by BANYULS in Plant Flavonoid Biosynthesis. *Science* **2003**, *299*, 396–399. [[CrossRef](#)]

42. Xie, D.-Y.; Jackson, L.; Cooper, J.; Ferreira, D.; Paiva, N. Molecular and Biochemical Analysis of Two cDNA Clones Encoding Dihydroflavonol-4-Reductase from *Medicago truncatula*. *Plant Physiol.* **2004**, *134*, 979–994. [[CrossRef](#)]
43. Losa, A.; Colombo, M.; Brambilla, V.; Colombo, L. Genetic interaction between AINTEGUMENTA (ANT) and the ovule identity genes SEEDSTICK (STK), SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2). *Sex. Plant Reprod.* **2010**, *23*, 115–121. [[CrossRef](#)] [[PubMed](#)]
44. Ezquer, I.; Mizzotti, C.; Nguema-Ona, E.; Gotté, M.; Beauzamy, L.; Viana, V.E.; Dubrulle, N.; Costa de Oliveira, A.; Caporali, E.; Koroney, A.; et al. The Developmental Regulator SEEDSTICK Controls Structural and Mechanical Properties of the Arabidopsis Seed Coat. *Plant Cell* **2016**, *28*, 2478–2492. [[CrossRef](#)]
45. Hsieh, W.-P.; Hsieh, H.-L.; Wu, S.-H. Arabidopsis bZIP16 transcription factor integrates light and hormone signaling pathways to regulate early seedling development. *Plant Cell* **2012**, *24*, 3997–4011. [[CrossRef](#)] [[PubMed](#)]
46. Hong, M.; Hu, K.; Tian, T.; Li, X.; Chen, L.; Zhang, Y.; Yi, B.; Wen, J.; Ma, C.; Shen, J.; et al. Transcriptomic Analysis of Seed Coats in Yellow-Seeded Brassica napus Reveals Novel Genes That Influence Proanthocyanidin Biosynthesis. *Front. Plant Sci.* **2017**, *8*, 1674. [[CrossRef](#)] [[PubMed](#)]
47. Wajid, B.; Serpedin, E. Do it yourself guide to genome assembly. *Brief. Funct. Genom.* **2014**, *15*, 1–9. [[CrossRef](#)] [[PubMed](#)]
48. Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-Cabrero, D.; Cervera, A.; McPherson, A.; Szczesniak, M.W.; Gaffney, D.J.; Elo, L.L.; Zhang, X.; et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* **2016**, *17*, 13. [[CrossRef](#)]
49. Chalhoub, B.; Denoeud, F.; Liu, S.; Parkin, I.; Tang, H.; Wang, X.; Chiquet, J.; Belcram, H.; Tong, C.; Samans, B. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. *Science (N. Y.)* **2014**, *345*, 950–953. [[CrossRef](#)]
50. Alexander, D.; Carrie, D.; Felix, S.; Jorg, D.; Chris, Z.; Sonali, J.; Philippe, B.; Mark, C.; Thomas, G. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)]
51. Laumont, C.M.; Daouda, T.; Laverdure, J.P.; Bonneil, A.; Caron-Lizotte, O.; Hardy, M.P.; Granados, D.P.; Durette, C.; Lemieux, S.; Thibault, P. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat. Commun.* **2016**, *7*, 10238. [[CrossRef](#)]
52. Ghosh, S.; Chan, C.-K.K. Analysis of RNA-Seq Data Using TopHat and Cufflinks. In *Plant Bioinformatics: Methods and Protocols*; Edwards, D., Ed.; Springer: New York, NY, USA, 2016; pp. 339–361.
53. Bevilacqua, V.; Pietroleonardo, N.; Giannino, E.I.; Stroppa, F.; Pesole, G.; Picardi, E. Clustering and Assembling Large Transcriptome Datasets by EasyCluster2. In *Emerging Intelligent Computing Technology and Applications*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 231–236.
54. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
55. Shihao, S.; Juw Won, P.; Zhixiang, L.; Lan, L.; Henry, M.D.; Yingnian, W.; Qing, Z.; Yi, X. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5593–5601. [[CrossRef](#)]
56. Li, D.; Dye, T.D. Power and stability properties of resampling-based multiple testing procedures with applications to gene oncology studies. *Comput. Math. Methods Med.* **2013**, *2013*, 610297. [[CrossRef](#)] [[PubMed](#)]

