



Article

# Genome-Wide Identification and Characterization of RNA/DNA Differences Associated with Drought Response in Wheat

Yan Pan <sup>1</sup>, Mengqi Li <sup>1</sup>, Jiaqian Huang <sup>1</sup>, Wenqiu Pan <sup>1</sup>, Tingrui Shi <sup>1</sup>, Qifan Guo <sup>1</sup>, Guang Yang <sup>1</sup>  
and Xiaojun Nie <sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory of Crop Stress Biology in Arid Areas, College of Agronomy, Northwest A&F University, Yangling 712100, China; py81212256@163.com (Y.P.); lmq13292903600@163.com (M.L.); Hjqqian1998@163.com (J.H.); wenqiu\_pan@nwfafu.edu.cn (W.P.); shitingrui@nwfafu.edu.cn (T.S.); gqf520521@163.com (Q.G.); drbiology@aliyuan.com (G.Y.)

<sup>2</sup> ICARDA-NWSUAF Joint Research Center, Yangling 712100, China

\* Correspondence: small@nwsuaf.edu.cn

**Abstract:** RNA/DNA difference (RDD) is a post-transcriptional RNA modification to enrich genetic information, widely involved in regulating diverse biological processes in eukaryotes. RDDs in the wheat nuclear genome, especially those associated with drought response or tolerance, were not well studied up to now. In this study, we investigated the RDDs related to drought response based on the RNA-seq data of drought-stressed and control samples in wheat. In total, 21,782 unique RDDs were identified, of which 265 were found to be drought-induced, representing the first drought-responsive RDD landscape in the wheat nuclear genome. The drought-responsive RDDs were located in 69 genes, of which 35 were differentially expressed under drought stress. Furthermore, the effects of RNA/DNA differences were investigated, showing that they could result in changes of RNA secondary structure, miRNA-target binding as well as protein conserved domains in the RDD-containing genes. In particular, the A to C mutation in TraesCS2A02G053100 (orthology to *OsRLCK*) led to the loss of *tae-miR9657b-5p* targeting, indicating that RNA/DNA difference might mediate miRNA to regulate the drought-response process. This study reported the first drought-responsive RDDs in the wheat nuclear genome. It sheds light on the roles of RDD in drought tolerance, and may also contribute to wheat genetic improvement based on epi-transcriptome methods.

**Keywords:** drought; miRNA targeting; RNA secondary structure; RNA/DNA difference; wheat



**Citation:** Pan, Y.; Li, M.; Huang, J.; Pan, W.; Shi, T.; Guo, Q.; Yang, G.; Nie, X. Genome-Wide Identification and Characterization of RNA/DNA Differences Associated with Drought Response in Wheat. *Int. J. Mol. Sci.* **2022**, *23*, 1405. <https://doi.org/10.3390/ijms23031405>

Academic Editor: Konrad Huppi

Received: 7 January 2022

Accepted: 24 January 2022

Published: 26 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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

Wheat (*Triticum aestivum* L.) is one of the most important cereal crops all over the world, providing the staple food source for about 30% of the global population and occupying approximately 20% of the world's cultivated lands [1]. Continuity of wheat production holds promise for ensuring global food security despite the challenges of population growth and global climate change [2–4]. It is estimated that wheat production should increase by 1.5% annually to meet the requirements of increasing population [5]. Drought is one of the most destructive environmental stresses that impair plant growth and development, reducing yields [6]. Recently, drought has become the most serious constraint to wheat production, causing about 5.5% yield loss every year [7,8]. Therefore, better understanding of the molecular mechanisms of drought response in wheat is of great significance for genetic improvement and breeding of drought-tolerant varieties.

RNA/DNA differences (RDDs) are phenomena of base insertion, deletion or modification, occurring when DNA is transcribed into RNA [9]. This widespread post-transcriptional modification mechanism contributes significantly to the diversity and plasticity of cellular RNA signatures, increasing proteomic diversity by modifying the sequence of primary transcripts that do not operate completely or partially [10,11]. Together with alternative splicing, RDD provides a crucial way to enrich genetic information and diversify transcripts,

playing an important role in regulating growth, development and stress response in eukaryotes [12]. ‘Substitution’ by simple base modification is the most common type of RDD, and is widely identified in plant organelle and higher eukaryotic nuclei genomes [13,14] as well as some viral sequences [15]. In mammals, the common type of RDD is A-to-I (G) (adenosine to inosine, guanosine), which is mainly catalyzed by double-stranded RNA-specific ADAR family proteins [16]. It is also called as A-to-G editing because the inosine in RNA is interpreted as guanosine (G) by the translation mechanism [17]. At the same time, A-to-I transformation independent of ADAR enzyme was also found in fungi [16]. However, ADAR-like enzymes have not been found in plants [18]. Another type, cytidine to uridine (C-to-U), is found to be catalyzed by APOBECs activation in humans [19], while the mechanism of the other remaining 10 modification types is still unclear [20,21]. In plants, RNA/DNA difference is generally found in organellar transcripts, and mainly regulated by the specific pentatripeptide repeat proteins encoded by the nuclear genome that deaminate cytidine to uridine [11,22,23]. There are two main types of PPR proteins: P and PLS [24]. The PLS group seems to be the most relevant class for RNA editing in plants with a conserved extended domain at the C-terminal and DYW motif [25–27]. The DYW motif encodes the active center of cytidine deaminase, which may be responsible for C-to-U modification [25].

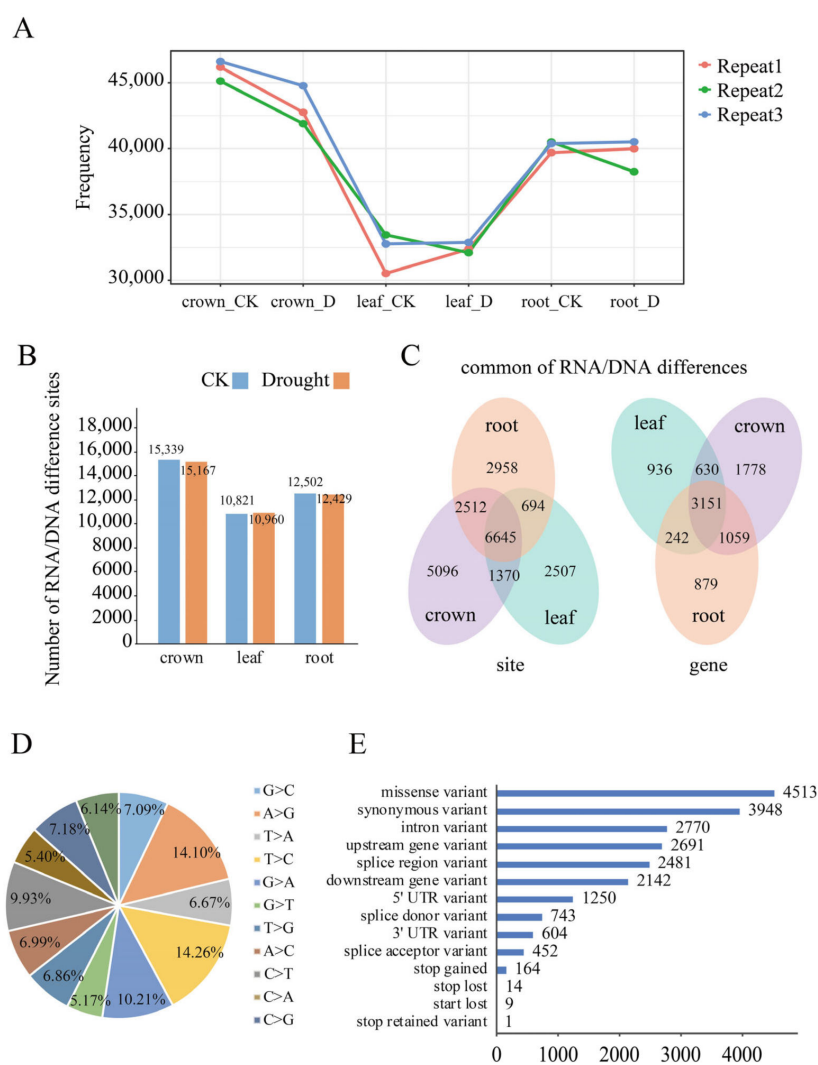
High-throughput RNA-seq technology makes it possible to investigate the transcript diversity and variation from the perspective of whole transcriptome level, which provides an efficient, unbiased, economical and comprehensive tool to identify RNA/DNA differences [28,29]. Extensive researches have utilized this method to identify the RNA editome and RNA/DNA difference landscape in yeast, humans and other model species, demonstrating the prevalence and significance of RNA/DNA differences [9]. However, the study of RNA/DNA difference in plants lags behind compared to animals and fungi, especially as genome-wide identification of RNA/DNA difference in plant nuclear genome was only conducted in Arabidopsis [30]. With the completion of a high-quality reference genome and a number of publicly available RNA-seq datasets, genome-wide identification of RNA/DNA differences in wheat nuclear genome become possible. Recently, Yang et al. reported the FHB-responsive RDDs in wheat nuclear genome through the RNA-seq dataset of the Fg-infected and mock-infected spike samples [31]. To obtain some insights on the RDDs related to drought stress, we used the 18 RNA-Seq datasets from 3 tissues under drought and control conditions of the cultivar Chinese spring (CS42) at the seedling stage to identify the drought-responsive RDDs here. Then, the effects of these RDDs on gene expression, RNA secondary structure, miRNA-target binding as well as protein conserved domain of the edited genes were also systematically investigated. This study aims to illuminate whether RDDs is occurred under drought stress and then to identify some drought-responsive RDD event, which will lay the foundation for better understanding the epitranscriptomic mechanism of drought response and tolerance in wheat.

## 2. Results

### 2.1. Identification of RNA/DNA Differences Based on RNA-Seq Data in Wheat

To identify potential RNA/DNA differences, we used the 18 RNA-Seq datasets of 3 tissues collected from the genotype CS42 seedlings under drought and control conditions, which is the same accession used to generate the reference genome. Based on the approach described in the Methods section, a total of 15,339, 10,821 and 12,502 RDDs were found in the control samples of crown, leaf and root respectively, while 15,167, 10,960, 12,429 RDDs were identified in drought-stressed samples of the crown, leaf and root, respectively (Figure 1A,B). It seems that the CK sample displayed slightly higher RDDs compared to the drought-treated sample in all of the 3 tissues. Meanwhile, for both the CK and drought condition, leaves had the lowest RDDs. These results indicated that RDDs were widely found in wheat nuclear genome. Furthermore, 21,782 unique RDDs were obtained through combining these identified RDDs, of which 8675 appeared in protein-coding genes (Table S1). Among them, 4205 genes had one RDD, followed by 1794, 1068 and 588 genes

with 2, 3 and 4 RDDs, and 1020 genes with more than 5 RDDs, respectively. Overall, 15,623, 11,216, 12,809 RDDs corresponding to 6618, 4959 and 5331 genes were found in crown, leaves and root respectively, of which 6645 RDDs in 3151 genes were commonly found in all of the three tissues (Figure 1C). For RNA/DNA difference type, all of the 12 conversion types were found, of which the conversion between C and T accounted for 24.19% of all sites, and conversion between A and G accounted for 24.31% respectively, which were the two canonical RNA/DNA difference types [31]. Additionally, C to G (7.18%) and G to C (7.09%), A to C (6.99%), T to G (6.86%), T to A (6.67%), A to T (6.14%), and C to A (5.40%) as well as G to T (5.17%) were also identified (Figure 1D). Gene annotation showed that 12,325, 2770, 1250 and 604 RDDs were located in the CDS, intron, 5'UTR and 3'UTR regions, respectively. Furthermore, 4513 and 3948 RDDs were annotated as missense variants and synonymous variants, accounting for 20.72% and 18.13% of RDDs, respectively, of which 189 RDDs could result in the premature stop or loss of the start/termination codons (Figure 1E).



**Figure 1.** Characterization of RNA/DNA differences in wheat nuclear genome using RNA-Seq data. (A) The number of variations identified in three biological replicates of control and drought conditions in the crown, leaf and root, respectively; (B) The numbers of unique RNA/DNA differences identified in crown, leaf and root, respectively; (C) The shared and unique RDD sites (Left) and RDD-containing genes (Right) found in crown, leaf and root tissues; (D) The variation types of the identified unique RDDs; (E) The frequency distribution of RDDs in the transcription regions. The y-axis represents different types of gene regions, and the x-axis represents the abundance of RDDs.

## 2.2. Identification of RDDs Associated with Drought Response

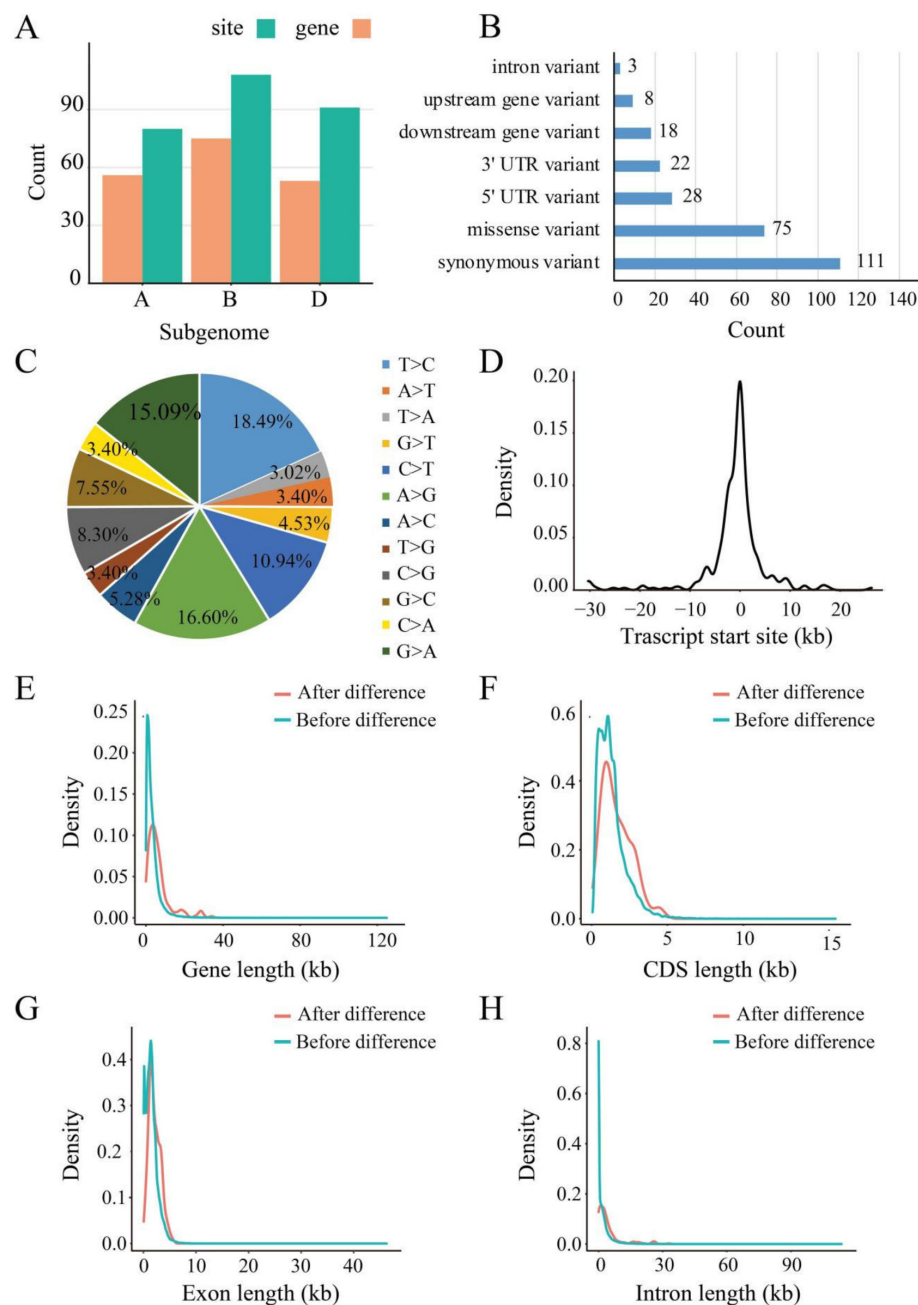
Based on the above identified RDDs, we further comprehensively screened the potential RDDs associated with drought stress through comparing the CK and stressed samples combined with IGV viewing. In total, 265 drought-responsive RDDs were obtained, of which 78, 97 and 90 RDDs were found in crown, leaves and root, respectively (Table S2). Surprisingly, drought-responsive RDD events displayed strong tissue specificity that none was shared by the three tissues. Among them, 76, 100 and 89 RDDs were located on subgenome A, B and D, corresponding to 54, 69 and 53 protein-coding genes, respectively (Figure 2A). Furthermore, 186 and 79 drought-responsive RDDs were located in CDS and UTR regions respectively, of which, 75 RDDs could cause missense variants (Figure 2B). These RDDs that could cause amino acid changes might play an important regulatory role in coping with drought stress response and tolerance. For difference type, there also all 12 types were found. Similar to the overall RDDs, the two canonical types conversion between C and T together with conversion between A and G were also the most abundant types (Figure 2C), similar to wild grapevine [32,33]. It is obvious that base transition events were higher than base transversion in these RDDs, although transversions should have twice higher occurrence frequency (transition/transversion ratio was 1.55), suggesting that these RDDs were not randomly occurring and related to the specific transcription regulation mechanism underlying the RND/DNA differences resulting from drought stress.

Thus, we further investigated the physical position of these drought-responsive RDDs. Results showed that they were mainly located about 10kb upstream or downstream of the translational start site (TSS) of their corresponding genes (Figure 2D); it is suggested that RDDs may impact on the expression of the corresponding genes. Compared to the genes that did not contain RDDs, the genes containing RDDs have significantly longer gene, exon, CDS and intron size (Figure 2E–H). Furthermore, the difference occurrence efficiency was also calculated based on the ratio of the number of difference reads to the total reads for each RDD site (Figure 3A). Results found that these RDDs under the drought condition displayed significantly higher difference efficiency than that of CK. The efficiency value of the drought-responsive RDDs ranged from 0.07 to 0.83, with most showing 0.1 to 0.2 (108 sites), followed by 0.2 to 0.3 (55 sites), 0.3 to 0.4 (41 sites), 0.4 to 0.5 (22 sites), 0.5 to 0.6 (13 sites), 0.6 to 0.7 (10 sites), 0 to 0.1 (9 sites), 0.7 to 0.8 (4 sites) and 0.8 to 0.9 (3 sites). The efficiency density of the three tissues is skewed to the left (Figure 3B), suggesting that all drought-responsive RDDs had relatively low difference efficiency in all tissues. Finally, we randomly selected one site in TraesCS3A02G045300 to perform experimental verification. Total RNA and DNA was extracted from the leaves. Results showed that the T to C mutation was not found in genomic DNA or the cDNA of CK samples, but appeared in the cDNA of drought stressed samples (Figure 3C), indicating that it is an actual drought-responsive RDD site.

## 2.3. Effects of RDDs on Gene Expression

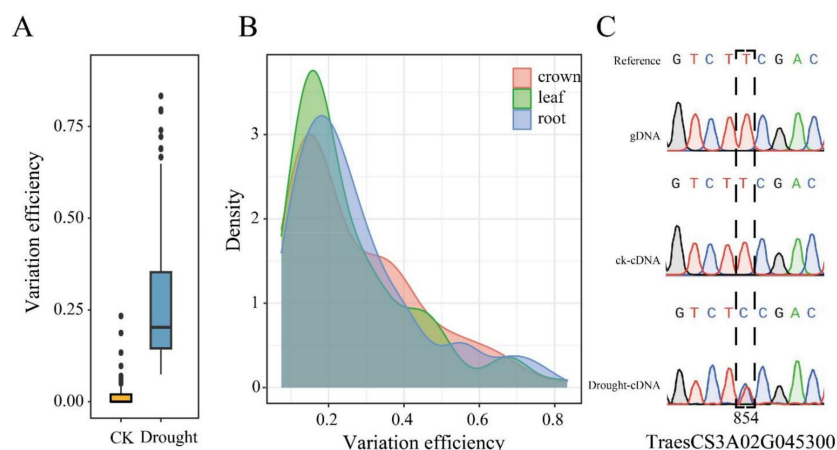
To explore the relationship between RDDs and gene expression, we further investigated the expression level of the genes containing drought-responsive RDDs under CK and drought conditions. In total, 35 unique genes were found to display differential expression patterns (Figure 4A), suggesting they might play a role in the regulation process of drought response. In detail, 11 (6 down-regulated and 5 up-regulated), 15 (9 down-regulated and 6 up-regulated) and 9 (8 down-regulated and 1 up-regulated) differential RDD-containing genes were found in root, leaf and crown, respectively. Among them, TraesCS2B02G038700 is annotated to encode a Chalcone synthase. Chalcone synthase is the key enzyme controlling flavonoid biosynthesis, which has been proven to monitor the drought stress tolerance [34]. TraesCS2B02G079100 is the orthology of *OsRBCS*, which is a drought stress related marker gene [35]. These results indicated that RDDs in these drought related genes might mediate their expression to regulate the drought response process. Furthermore, functional enrichment found that they are significantly enriched in stress response related terms, such as ‘defense response’, ‘response to oxidative stress’

and ‘response to desiccation’. In addition, they are also enriched in terms related to the structure of DNA or RNA, including ‘mRNA binding’, ‘DNA binding’. In KEGG pathway, they are enriched in the biosynthesis of secondary metabolites, flavonoids biosynthesis, and zeatin biosynthesis, all of which were also related to the drought response in plants (Figure 4B) (Table S3).

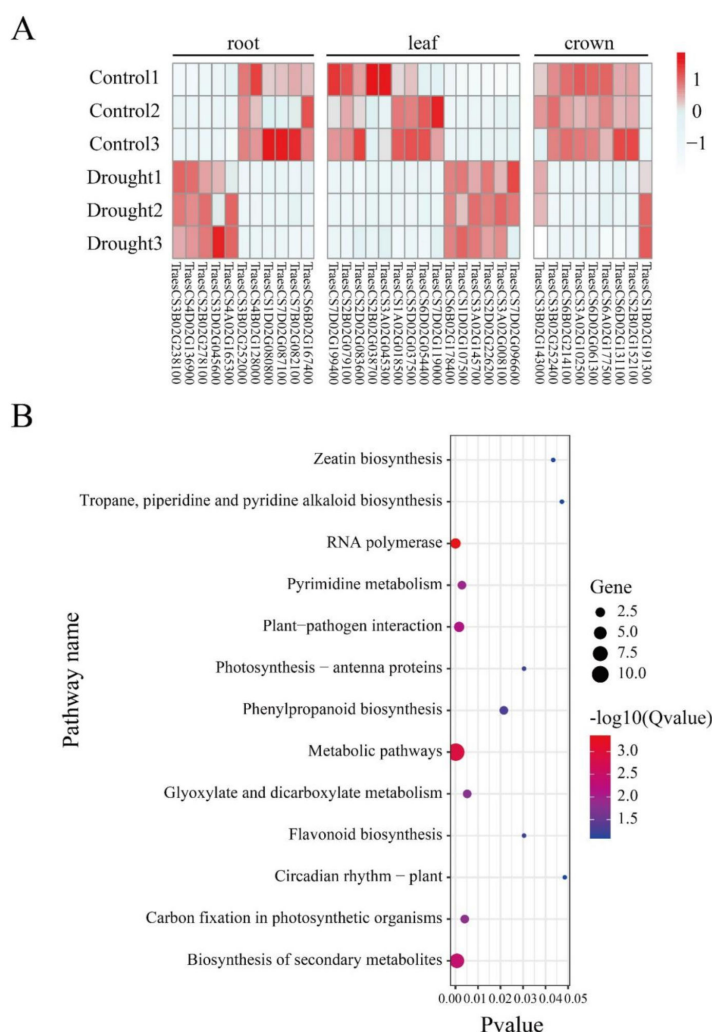


**Figure 2.** Characterization of the drought-responsive RDDs. (A) The numbers of drought-responsive RDD and related-genes in A, B, D subgenome, respectively; (B) The distribution of drought-responsive RDDs in the transcription regions. The  $y$ -axis represents different types of regions, and the  $x$ -axis represents the abundance of RNA/DNA difference sites; (C) The variation types of the identified drought-responsive RDDs; (D) The distance of the distribution of drought-responsive RDD site and the TSS (transcription start site) in the RDD-containing genes; (E–H) Comparison of gene length, CDS length, exon length, and intron length of gene without RDD and with RDD.





**Figure 3.** Variation efficiency and validation of these identified RDDs. (A) Comparison of the variation efficiency of RDDs between CK and drought conditions; (B) the distribution of variation efficiency in 3 different tissues; (C) Validation of the randomly selected drought-responsive RDD in TraesCS3A02G045300 through RT-PCR combined with Sanger sequencing.



**Figure 4.** Differential expression and functional enrichment of the genes occurring in drought-responsive RDD. (A) The differential expression of RDD-containing genes between CK and drought conditions in roots, leaf and crown tissues; (B) KEGG enrichment analysis of these differentially expressed RDD-containing genes.

#### 2.4. Effect of RDDs on RNA Secondary Structure and MiRNA Targeting

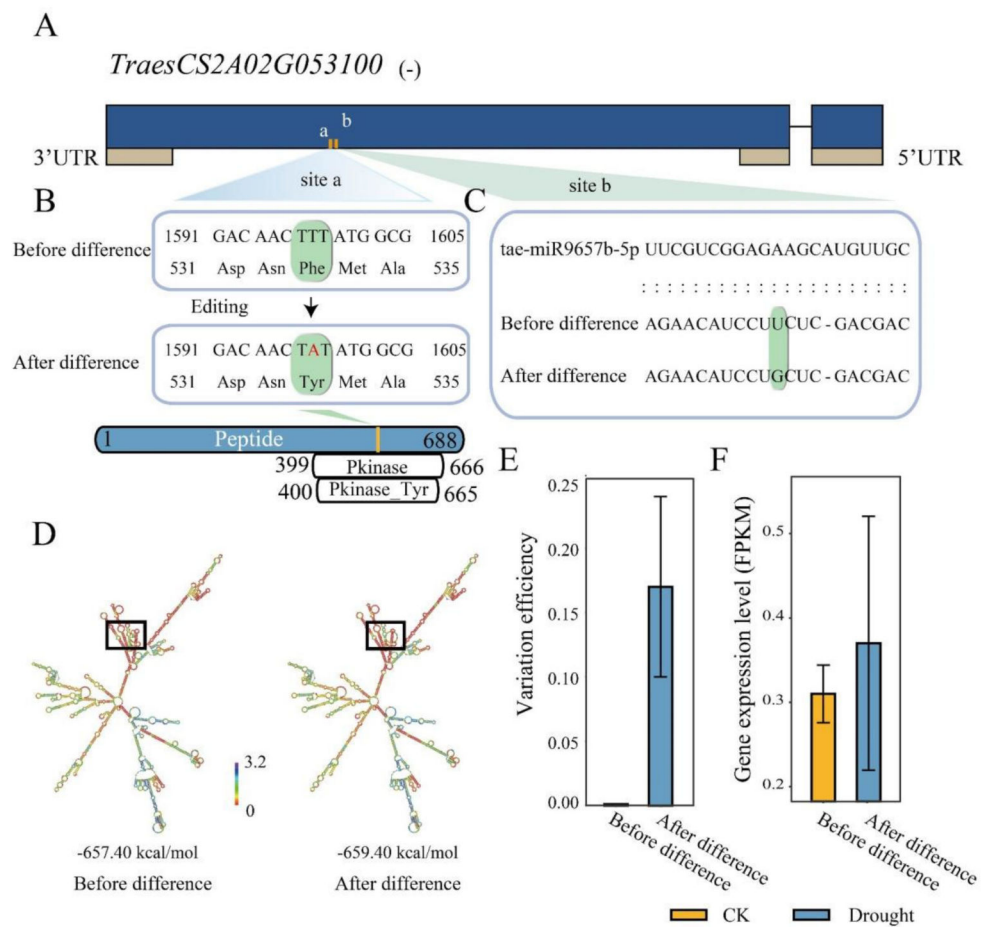
Secondary structural transformation, as a binary switch activated by cellular signals, is a general mechanism in gene regulation networks [36]. We further investigated the effects of RDDs on RNA secondary structure and found that 171 RDDs in 117 genes can lead to the changes of their RNA secondary structure (Table S4). Among them, the minimum free energy (MFE) of 61 genes increased after RDD occurred, while that of the other remaining 56 sites decreased (Table S4). As we know, MFE can be used to measure the stability of the RNA molecule, and a structure with a low MFE value is more stable [37,38]. According to our prediction, drought responsive RDD could reduce the stability of 61 genes and also increase the stability of the other 56 genes. The secondary structure changes of these genes will impact on achieving their function, indicating that the RDDs might be involved in regulating drought response through influencing the structural stability of the drought-related genes. Further study on the specific role of RDD in regulating RNA secondary structure will contribute to better understanding of the genetic basis of crop drought resistance.

miRNA is a class of small non-coding RNAs that often bind to complementary sequences of the mRNA targets to silence or weaken their expression [30]. RDD generally caused variations in mRNA sequence, which could make the transcripts gain or lose miRNA binding [39]. To understand the effects of drought responsive RDDs on miRNA targeting, we further investigated miRNA-mRNA targeting pairs using the genes with or without RDD mutations as target genes. Results showed that there were 160 genes that were potentially targeted by 81 unique high-confidence mature miRNAs of wheat, containing 236 RDDs (Table S5), of which 3 RDDs occurring in 3 genes were detected to result in the changing of miRNA targets compared to the initial genes without RDD. We found two RDD sites (chr2A\_21252947: A to T; chr2A\_21252961: A to C) occurring in TraesCS2A02G053100 (Figure 5A): the T to A site changed the amino acid from Phe to Tyr (Figure 5B), and the T to G site made it lose the binding of taе-miR9657b-5p (Figure 5C). Meanwhile, the RNA secondary structure of this gene was also changed by the RDD sites (Figure 5D), that the MFE value changed from  $-657.40$  to  $-659.40$  kcal/mol. Furthermore, the difference efficiency of T to G in TraesCS2A02G053100 is 0.18 under the drought condition while not occurring in CK with the variation efficiency of 0 (Figure 5E). The expression level of this gene was also slightly higher under the drought condition compared to CK (Figure 5F). These results indicated that RDD could precisely regulate the expression level of the target genes through mediating miRNA targeting with variable variation efficiency. In addition, orthologs analysis found that TraesCS2A02G053100 is the ortholog of *OsRLCK*. *OsRLCK* is reported as the key candidate associated with QTL for drought stress related traits [40], suggesting that TraesCS2A02G053100 could also have potential function in response to drought stress. The RDD occurring in this gene might play an indispensable role in regulating the sophisticated process of drought response that this gene was involved in.

#### 2.5. Effects of RDDs on Protein Conserved Domain

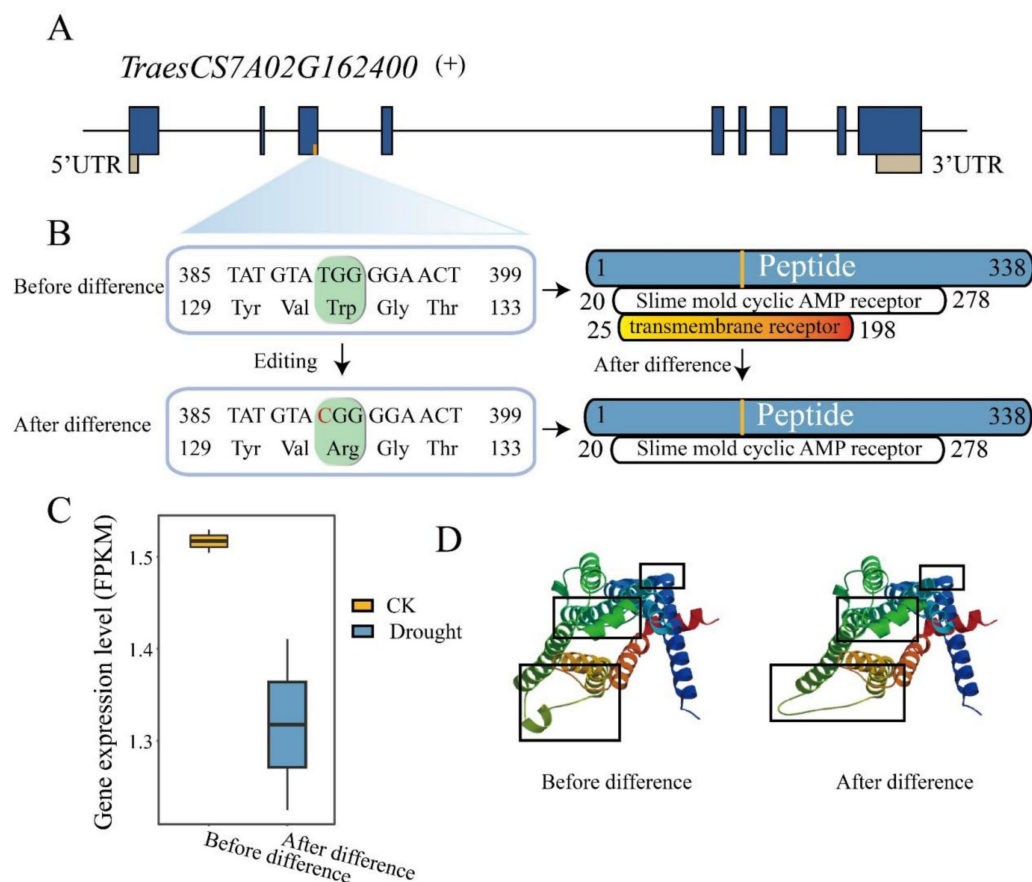
RNA/DNA difference is a post-transcriptional modification mechanism to increase proteomic diversity by modifying protein composition, protein structure and binding ability [41]. To understand the effect of drought-responsive RDD on the encoded protein structure, we investigated and compared the conserved domain organization of the genes with missense variant and stop gained due to RDD before and after RDD occurred. The results showed that 2 genes were impacted by RDDs to change their protein domain organization (Table S6), suggesting that RDDs might have effects on protein functions. One is TraesCS6D02G181700, which is the orthology of *AtABA2*, having two RDD sites (chr6D\_208258669; chr6D\_208258671) and site chr6D\_208258671, changing the amino acid from Val to Ala to provide an wpimerase domain after RDD occurred. It is reported that *AtABA2* encoded xanthin dehydrogenase, which is a key enzyme for abscisic acid (ABA) biosynthesis, to play the crucial role in regulating ABA levels in Arabidopsis [42,43]. The

other gene is TraesCS7A02G162400, which is the orthology of G-protein coupled receptor 1 (*GCR1*) in Arabidopsis. A RDD site T to C (chr7A\_118456566) in TraesCS7A02G162400 resulted in loss of the transmembrane receptor domain although the slime mold cyclic AMP receptor domain was present (Figure 6A,B). Previous studies have demonstrated that *GCR1* was an important regulator involved in stress signal transduction, such as drought stress, ABA response and regulation of stomatal aperture in Arabidopsis [44,45]. Furthermore, we found that the expression level of this gene was also down-regulated after RDD occurred (Figure 6C), and the protein three-dimensional (3D) structure predictions also displayed differences before and after RDD (Figure 6D). These results indicated the drought-responsive RDD could also impact on the protein domain of the drought-related genes and act as the regulator in the biological process of drought response and tolerance.



**Figure 5.** RNA/DNA difference functions on the miRNA targeting in *TraesCS2A02G053100*. (A) Two RNA/DNA difference sites (a, b) were found in *TraesCS2A02G053100*, which is annotated in the reverse chain of the genome; (B) The RDD site A to T (site a) caused amino acid Phe changed into Tyr; (C) The RDD site A to C (site b) leads to the loss of miR9657-5p targeting; (D) Change of the RNA 2D structure before and after RDD variations; (E) Variation efficiency of the A to C site (site b); (F) Expression levels of *TraesCS2A02G053100* before and after variation.





**Figure 6.** RNA/DNA difference impacts on the protein conserved domains and 3D structure in *TraesCS7A02G162400*. (A,B) A T to C variation was identified in the third exon of *TraesCS7A02G162400*, which caused Trp to Arg change, and lost a protein conserved domain; (C) Expression levels of *TraesCS7A02G162400* before and after variation; (D) Changes of protein 3D structure before and after difference.

### 3. Discussion

Plants have evolved diverse regulatory mechanisms to deal with environmental stresses at both the transcriptional and post-transcriptional levels [46,47]. From the perspective of post-transcriptional regulation, little is known about the role and function of RNA/DNA difference in response to drought although some studies have been conducted to identify alternative splicing events associated with drought stress [48,49]. Here, we conducted the first global survey of drought-related RDDs in wheat and 265 drought-responsive RDDs in 176 genes were obtained. Furthermore, we validated one drought-responsive RDD using the RT-PCR combined with Sanger sequencing. This study not only proved that RDD events actually occurred in the nuclear genome in wheat, but also demonstrated that they were involved in regulating drought response. Although we cannot assure that all of the remaining 264 RDD sites were in actual existence due to possible false positive results, we used the most rigorous and reliable current approaches to obtain reliable results. Firstly, the RNA-Seq dataset from the same accession used for generating the reference genome was adopted to identify RDDs, which completely ruled out the interference of genotype-specific mutation and polymorphism. Secondly, three independent biological replications were adopted and only the RDD found in all of the 3 replications was retained, which could excluded sequencing errors or other random mistakes. Thirdly, comparison of the control and drought-stressed sample combined with manual verification through IGV checking were used to obtain the high-confidence drought-responsive RDDs, which can remove the mapping errors. At the same time, the wheat genome is one of

the most complex genomes in plants, with A, B and D subgenomes as well as more than 80% repetitive elements, which could result in false positive mapping when RNA-Seq is mapped on the reference genome causing false positive RDD results [50]. Thus, the identified drought-responsive RDD reported here were reliable within the current level of knowledge and technology.

Furthermore, we found there is no significant correlation of RDD with its gene expression. Only a small part of genes with drought-responsive RDDs displayed differential expression under drought stress compared to the control, indicating that RDD exerted slight influence on gene expression. Then, we detected the effects of RDDs on the RNA secondary structure. Out of 265 drought-responsive RDDs, 171 RDDs lead to changes of RNA secondary structure in 117 RDD-containing genes, of which 61 genes reduced the stability and 56 genes increased the stability due to RDDs, suggesting that RDDs impacted significantly on the RNA composition and then affected the RNA structure. It is reported that post-transcriptional gene expression mediated by small non-coding RNAs (such as microRNAs) and RDD events regulatory processes are connected at a fundamental level [51]. In humans, the functional relationship between RDDs and miRNA mediated post transcriptional gene silencing has been revealed [51]. RDD can result in the gain or loss of miRNA target sites in the RDD-containing genes [30]. We further predicted and compared the miRNA targets of all RDD-containing genes affected by RDDs, and found that 3 RDD events caused the change of miRNA targeting. In particular, T to G difference in TraesCS2A02G053100 caused it to lose the binding of taе-miR9657b-5p, and then increase its expression under the drought condition. Orthologous gene analysis found that TraesCS2A02G053100 might be a drought-related gene. It is noticed that RDD occurs alongside efficiency, suggesting that it can regulate the expression level of RDD-containing gene with a very precise method; this may quantify the difference efficiency to produce the quantitative transcript, and then mediate the miRNA to participate in the drought response process. Further study of the role of RDD in mediating miRNA targeting could elucidate a novel regulatory mechanism or network underlying drought response and tolerance in wheat [51].

Generally, one important effect of RDDs is to increase the diversity of the proteome. Here, 75 out of 265 drought-responsive RDDs were found to cause the missense variants, which contributed to enriching proteomic diversity. The effects of RDD on the protein conserved domain and 3D structures were also investigated. As we know, structure is the basis of protein function. The change of protein domain and 3D structure caused by RDDs might also affect the protein's function. In this study, we found that a drought-responsive RDD (T to C) in TraesCS7A02G162400, which is the orthology of *AtGCR1*, resulted in loss of the transmembrane receptor domain and also changed its 3D structure. It is reported that *GCR1* acts as a negative regulator of GPA1-mediated ABA response in Arabidopsis guard cells, and loss-of-function mutants of *AtGCR1* showed resistance to drought stress and activated the high expression of some known drought and ABA regulatory genes under drought stress [45]. The drought-responsive RDD in TraesCS7A02G162400 might regulate this gene to be involved in the drought stress response through functions on the protein domain and 3D structure. Further functional studies of this RDD are needed to reveal its role in regulating drought stress, enriching the epigenetic mechanism of drought response and tolerance in wheat.

## 4. Materials and Methods

### 4.1. RNA-Seq Data and Mapping

The RNA-Seq dataset used in this study was downloaded from Sequence Read Archive database with the accession no. SRP098756 [52]. Raw data was quality controlled by FastQC (version 0.11.8) and Trimomatic (version 0.39) to remove adapter reads, low-quality reads, or contamination. The obtained high-quality reads were aligned to wheat reference genome (IWGSC RefSeq version 1.1) using STAR tool (version 2.7.6). The alignment file was subsequently analyzed using StringTie (version 1.3.5) and subsequently the gtf file was

combined using the merge function in StringTie of each sample. In addition, we quantified the read coverage of each gene by featurecount (version 0.11.2). Differential gene expression analysis was performed by using DESeq2 tool with the adjusted P value as less than 0.05 and  $|\log_2FC| > 1$ .

#### 4.2. Identification of RNA/DNA Difference Sites

These RNA-Seq reads were aligned to the wheat reference genome and removed from the same location reads using MarkDuplicates tool in Picard [53]. Then, the SplitNCigarReads tool in GATK (Genome Analysis Toolkit) software was used to separate the reads on the exons, remove the N error bases, and remove the reads in the intron region. Then, the AddOrReplaceReadGroups tool in GATK software was used to assign all read operations in a file to a new read group. Finally, the HaplotypeCaller tool in GATK software was used to call DNA/RNA differences. Subsequently, DNA/RNA differences of each sample were obtained and stored in the raw VCF file for further analysis to obtain the potential RDDs as follows:

- (1) GATK VariantFiltration tool was used to correct the systematic errors of the sequencing platform and software, and the initial filtering parameters – Filter “FS > 30.0”, – Filter “QD < 2.0” were selected;
- (2) The site was mapped by more than 10 reads with reference read >2 and mutation read >3 retained;
- (3) In order to improve the accuracy, only the DNA/RNA difference sites that appeared in all of the three replicates were retained;
- (4) the RDDs between control and drought-treated samples were compared to remove the same sites to obtain the potential drought-responsive RDDs;
- (5) manually verify each bam file to assure the RDD’s actual occurrence through Integrative Genomics Viewer (IGV) tool.

Through these filtering, the high-confidence RDDs associated with drought were finally obtained. Then, we used the SnpEff tool (version 3.6) and the annotation file downloaded from the Ensembleplants database [54] to annotate the filtered VCF file.

#### 4.3. Validation of RNA/DNA Difference Site Using RT-PCR

For experiment verification, the seedlings of genotype CS42 grown under normal and drought (19.2% PEG) conditions were collected to isolate RNA using the RNA Easy Fast Plant Tissue Kit (Tiangen, Beijing, China). Then, total RNA was used to synthesize cDNAs using RT Master Mix Perfect Real-Time kit (Takara, Dalian, China) according to the manufacturer’s instruction. We randomly selected one RDD (T to C) in TraesCS3A02G045300 for validation by RT-PCR analysis using the forward primer and reverse primer of GAA-GATCTGGTTTCCATGGGACT and GCAAATCATGCAGTGGAATCAAGC, respectively.

#### 4.4. RNA Secondary Structure and MiRNA Target Analysis

The RNA secondary structure of the RDD-containing gene was predicted by the RNAfold online tool [55] using the default parameters. To determine whether RDDs affected miRNA targeting, all of the RDD-containing genes were searched against the published 119 wheat miRNAs in the miRBase database using psRNATarget tools [56]. Schema V2 (2017 release) model was used to score the possibility of miRNA targeting on the gene, and the result with the lowest expectation was selected as the optimal prediction. Then, the effects of RDDs on miRNA binding were obtained through comparison of the prediction results of whether RDD occurred or not.

#### 4.5. Conserved Domain Prediction and Protein Structure Analysis

The PFAM database (version 33.0) was used to search protein domains based on HMMER3 (version v3.1.1) for RDD-containing genes using protein sequences as input with the E-value  $< 1 \times 10^{-5}$ . The 3D structure was predicted using the homologous modeling

method in SWISS-MODEL database [57]. The model with the highest degree of coincidence with the target protein and more than 30% was adopted.

## 5. Conclusions

Here, we identified the first drought-responsive RDD landscape in wheat, proving that RDD events actually occurred in the wheat nuclear genome and were also involved in regulating drought response. Furthermore, the effects of these drought-responsive RDDs were systematically investigated, showing that these RDDs could exert diverse effects on the drought-related genes to mediate drought response and tolerance. They could impact on the expression, RNA secondary structure, miRNA targeting as well as protein domains and 3D structure of the RDD-containing genes. The identified drought-responsive RDDs and their RDD-containing genes provide important resources for mining new genes associated with drought stress. This study lays the foundation for further functional studies to better understand the epigenetic mechanisms underlying drought stress, and also pave the way to identify RDDs in the nuclear genome based on RNA-Seq data in wheat and beyond.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23031405/s1>.

**Author Contributions:** X.N. designed and supervised the project. Y.P., M.L. and J.H. collected and generated the data, and performed analysis. W.P. and T.S. performed RT-PCR analysis and contributed to data analysis. Q.G. and G.Y. helped to prepare the figures and tables. Y.P. prepared the draft manuscript. X.N. reviewed and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was mainly funded by the National Natural Science Foundation of China (Grant No. 31771778 and 31971885), and partially supported by the Key Research and Development Program of Shaanxi Province, China (Grant No. 2019NY-014).

**Data Availability Statement:** The data that supports the findings of this study are available in the supplementary material of this article.

**Acknowledgments:** We are grateful to Licao Cui and Ruoyu Zhang for their help on data retrieval and we appreciate the High-Performance Computing center of Northwest A&F University for providing computational resources.

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

## References

1. Shiferaw, B.; Smale, M.; Braun, H.J.; Duveiller, E.; Reynolds, M.; Muricho, G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* **2013**, *5*, 291–317. [[CrossRef](#)]
2. FAO Statistics. Available online: <http://www.fao.org/unfao/procurement/statistics-from-2010-2020/statistics-2019/en/> (accessed on 4 June 2021).
3. Asseng, S.; Ewert, F.; Martre, P.; Rötter, R.P.; Lobell, D.B.; Cammarano, D.; Kimball, B.A.; Ottman, M.J.; Wall, G.; White, J.W.; et al. Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* **2015**, *5*, 143–147. [[CrossRef](#)]
4. Zampieri, M.; Ceglar, A.; Dentener, F.; Toreti, A. Wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales. *Environ. Res. Lett.* **2017**, *12*, 064008. [[CrossRef](#)]
5. Stratonovitch, P.; Semenov, M.A. Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *J. Exp. Bot.* **2015**, *66*, 3599–3609. [[CrossRef](#)] [[PubMed](#)]
6. Lipiec, J.; Doussan, C.; Nosalewicz, A.; Kondracka, K. Effect of drought and heat stresses on plant growth and yield: A review. *Int. Agrophys.* **2013**, *27*, 463–477. [[CrossRef](#)]
7. Basu, S.; Ramegowda, V.; Kumar, A.; Pereira, A. Plant adaptation to drought stress. *F1000Research* **2016**, *5*. [[CrossRef](#)]
8. Fahad, S.; Bajwa, A.A.; Nazir, U.; Anjum, S.A.; Farooq, A.; Zohaib, A.; Sadia, S.; Nasim, W.; Adkins, S.; Saud, S.; et al. Crop production under drought and heat stress: Plant responses and management options. *Front. Plant Sci.* **2017**, *8*, 1147. [[CrossRef](#)]
9. Peng, Z.; Cheng, Y.; Tan, B.C.M.; Kang, L.; Tian, Z.; Zhu, Y.; Zhang, W.; Liang, Y.; Hu, X.; Tan, X.; et al. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* **2012**, *30*, 253–260. [[CrossRef](#)]
10. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 83–96. [[CrossRef](#)]
11. Gray, M. Diversity and evolution of mitochondrial RNA editing systems. *IUBMB Life* **2003**, *55*, 227–233. [[CrossRef](#)]



12. Bian, Z.; Ni, Y.; Xu, J.R.; Liu, H. A-to-I mRNA editing in fungi: Occurrence, function, and evolution. *Cell. Mol. Life Sci.* **2019**, *76*, 329–340. [[CrossRef](#)] [[PubMed](#)]
13. Qiu, Y.L.; Li, L.; Wang, B.; Chen, Z.; Dombrowska, O.; Lee, J.; Kent, L.; Li, R.; Jobson, R.W.; Hendry, T.A.; et al. A nonflowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* **2007**, *168*, 691–708. [[CrossRef](#)]
14. Takenaka, M.; Verbitskiy, D.; van der Merwe, J.A.; Zehrmann, A.; Brennicke, A. The process of RNA editing in plant mitochondria. *Mitochondrion* **2008**, *8*, 35–46. [[CrossRef](#)] [[PubMed](#)]
15. Wang, L.; Sun, Y.; Song, X.; Wang, Z.; Zhang, Y.; Zhao, Y.; Peng, X.; Zhang, X.; Li, C.; Gao, C.J.; et al. Hepatitis B virus evades immune recognition via RNA adenosine deaminase ADAR1-mediated viral RNA editing in hepatocytes. *Cell. Mol. Immunol.* **2021**, *18*, 1871–1882. [[CrossRef](#)]
16. Savva, Y.A.; Rieder, L.E.; Reenan, R.A. The ADAR protein family. *Genome Biol.* **2012**, *13*, 252. [[CrossRef](#)]
17. Nishikura, K. Editor meets silencer: Crosstalk between RNA editing and RNA interference. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 919–931. [[CrossRef](#)]
18. Jin, Y.; Zhang, W.; Li, Q. Origins and evolution of ADAR-mediated RNA editing. *IUBMB Life* **2009**, *61*, 572–578. [[CrossRef](#)]
19. Salter, J.D.; Bennett, R.P.; Smith, H.C. The APOBEC protein family: United by structure, divergent in function. *Trends Biochem. Sci.* **2016**, *41*, 578–594. [[CrossRef](#)]
20. Li, M.; Wang, I.X.; Li, Y.; Bruzel, A.; Richards, A.L.; Toung, J.M.; Cheung, V.G. Widespread RNA and DNA sequence differences in the human transcriptome. *Science* **2011**, *333*, 53–58. [[CrossRef](#)]
21. Liscovitch-Brauer, N.; Alon, S.; Porath, H.T.; Elstein, B.; Unger, R.; Ziv, T.; Admon, A.; Levanon, E.Y.; Rosenthal, J.J.; Eisenberg, E. Trade-off between transcriptome plasticity and genome evolution in cephalopods. *Cell* **2017**, *169*, 191–202. [[CrossRef](#)]
22. Qulsum, U.; Azad, M.T.A.; Tsukahara, T. Analysis of tissue-specific RNA editing events of genes involved in RNA editing in *Arabidopsis thaliana*. *J. Plant Biol.* **2019**, *62*, 351–358. [[CrossRef](#)]
23. Small, I.D.; Schallenberg-Rüdinger, M.; Takenaka, M.; Mireau, H.; Ostersetzer-Biran, O. Plant organellar RNA editing: What 30 years of research has revealed. *Plant J.* **2020**, *101*, 1040–1056. [[CrossRef](#)]
24. Wang, X.; An, Y.; Xu, P.; Xiao, J. Functioning of PPR Proteins in organelle RNA metabolism and chloroplast biogenesis. *Front. Plant Sci.* **2021**, *12*, 1. [[CrossRef](#)] [[PubMed](#)]
25. Liu, X.Y.; Jiang, R.C.; Wang, Y.; Tang, J.J.; Sun, F.; Yang, Y.Z.; Tan, B.C. ZmPPR26, a DYW-type pentatricopeptide repeat protein, is required for C-to-U RNA editing at atpA-1148 in maize chloroplasts. *J. Exp. Bot.* **2021**, *72*, 4809–4821. [[CrossRef](#)]
26. Ichinose, M.; Sugita, M. The DYW domains of pentatricopeptide repeat RNA editing factors contribute to discriminate target and non-target editing sites. *Plant Cell Physiol.* **2018**, *59*, 1652–1659. [[CrossRef](#)]
27. Malbert, B.; Burger, M.; Lopez-Obando, M.; Baudry, K.; Launay-Avon, A.; Härtel, B.; Verbitskiy, D.; Jörg, A.; Berthomé, R.; Lurin, C.; et al. The Analysis of the Editing Defects in the *dyw2* Mutant Provides New Clues for the Prediction of RNA Targets of *Arabidopsis E+*-Class PPR Proteins. *Plants* **2020**, *9*, 280. [[CrossRef](#)] [[PubMed](#)]
28. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63. [[CrossRef](#)] [[PubMed](#)]
29. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)]
30. Meng, Y.; Chen, D.; Jin, Y.; Mao, C.; Wu, P.; Chen, M. RNA editing of nuclear transcripts in *Arabidopsis thaliana*. *BMC Genom.* **2010**, *11* (Suppl. S4), S12. [[CrossRef](#)]
31. Yang, G.; Pan, Y.; Zhang, R.; Huang, J.; Pan, W.; Cui, L.; Song, W.; Nie, X. Genome-wide identification and characterization of DNA/RNA differences associated with *Fusarium graminearum* infection in wheat. *bioRxiv* **2021**. [[CrossRef](#)]
32. Ramos, M.J.; Faisca-Silva, D.; Coito, J.L.; Cunha, J.; Silva, H.G.; Viegas, W.; Costa, M.M.R.; Amâncio, S.; Rocheta, M. RNA editing in inflorescences of wild grapevine unveils association to sex and development. *bioRxiv* **2020**. [[CrossRef](#)]
33. Ichinose, M.; Sugita, M. RNA editing and its molecular mechanism in plant organelles. *Genes* **2017**, *8*, 5. [[CrossRef](#)] [[PubMed](#)]
34. Hu, B.; Yao, H.; Gao, Y.; Wang, R.; Jin, L. Overexpression of chalcone synthase gene improves flavonoid accumulation and drought tolerance in tobacco. *Preprints* **2019**. [[CrossRef](#)]
35. Lee, D.K.; Jung, H.; Jang, G.; Jeong, J.S.; Kim, Y.S.; Ha, S.H.; Choi, Y.D.; Kim, J.K. Overexpression of the OsERF71 Transcription Factor Alters Rice Root Structure and Drought Resistance. *Plant Physiol.* **2016**, *172*, 575–588. [[CrossRef](#)] [[PubMed](#)]
36. Dethoff, E.A.; Chugh, J.; Mustoe, A.M.; Al-Hashimi, H.M. Functional complexity and regulation through RNA dynamics. *Nature* **2012**, *482*, 322–330. [[CrossRef](#)]
37. Andronescu, M.; Fejes, A.P.; Hutter, F.; Hoos, H.H.; Condon, A.J. A New Algorithm for RNA Secondary Structure Design. *J. Mol. Biol.* **2004**, *336*, 607–624. [[CrossRef](#)]
38. Yuan, L.; Zhang, S. Finding stable local optimal RNA secondary structures. *Bioinformatics* **2011**, *27*, 2994–3001.
39. Picardi, E.; Horner, D.S.; Chiara, M.; Schiavon, R.; Valle, G.; Pesole, G. Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. *Nucleic Acids Res.* **2010**, *38*, 4755–4767. [[CrossRef](#)]
40. Vij, S.; Giri, J.; Dansana, P.K.; Kapoor, S.; Tyagi, A.K. The receptor-like cytoplasmic kinase (OsRLCK) gene family in rice: Organization, phylogenetic relationship, and expression during development and stress. *Mol. Plant* **2008**, *1*, 732–750. [[CrossRef](#)]
41. Gott, J.M. Expanding genome capacity via RNA editing. *Comptes Rendus Biol.* **2003**, *326*, 901–908. [[CrossRef](#)]



42. González-Guzmán, M.; Apostolova, N.; Bellés, J.M.; Barrero, J.M.; Piqueras, P.; Ponce, M.R.; Micol, J.L.; Serrano, R.; Rodríguez, P.L. The short-chain alcohol dehydrogenase ABA2 catalyzes the conversion of xanthoxin to abscisic aldehyde. *Plant Cell* **2002**, *14*, 1833–1846. [[CrossRef](#)] [[PubMed](#)]
43. Endo, A.; Nelson, K.M.; Thoms, K.; Abrams, S.R.; Nambara, E.; Sato, Y.J. Functional characterization of xanthoxin dehydrogenase in rice. *J. Plant Physiol.* **2014**, *171*, 1231–1240. [[CrossRef](#)] [[PubMed](#)]
44. Pandey, A.; Sharma, M.; Pandey, G.K. Small and large g proteins in biotic and abiotic stress responses in plants. In *Elucidation of Abiotic Stress Signaling in Plants: Functional Genomics Perspectives*; Pandey, G.K., Ed.; Springer: New York, NY, USA, 2015; pp. 231–270.
45. Pandey, S.; Assmann, S.M. The Arabidopsis putative G protein–coupled receptor GCR1 interacts with the G protein  $\alpha$  subunit GPA1 and regulates abscisic acid signaling. *Plant Cell* **2004**, *16*, 1616–1632. [[CrossRef](#)] [[PubMed](#)]
46. Liu, Z.; Qin, J.; Tian, X.; Xu, S.; Wang, Y.; Li, H.; Wang, X.; Peng, H.; Yao, Y.; Hu, Z.; et al. Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* **2018**, *16*, 714–726. [[CrossRef](#)] [[PubMed](#)]
47. Thatcher, S.R.; Danilevskaya, O.N.; Meng, X.; Beatty, M.; Zastrow-Hayes, G.; Harris, C.; Van Allen, B.; Habben, J.; Li, B. Genome-wide analysis of alternative splicing during development and drought stress in maize. *Plant Physiol.* **2016**, *170*, 586–599. [[CrossRef](#)] [[PubMed](#)]
48. Wei, H.; Lou, Q.; Xu, K.; Yan, M.; Xia, H.; Ma, X.; Yu, X.; Luo, L. Alternative splicing complexity contributes to genetic improvement of drought resistance in the rice maintainer HuHan2B. *Sci. Rep.* **2017**, *7*, 11686. [[CrossRef](#)]
49. Ding, Y.; Wang, Y.; Qiu, C.; Qian, W.; Xie, H.; Ding, Z. Alternative splicing in tea plants was extensively triggered by drought, heat and their combined stresses. *PeerJ* **2020**, *8*, e8258. [[CrossRef](#)]
50. Zhu, T.; Wang, L.; Rimbart, H.; Rodriguez, J.C.; Deal, K.R.; De Oliveira, R.; Choulet, F.; Keeble-Gagnère, G.; Tibbits, J.; Rogers, J.; et al. Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *Plant J.* **2021**, *107*, 303–314. [[CrossRef](#)]
51. Borchert, G.M.; Gilmore, B.L.; Spengler, R.M.; Xing, Y.; Lanier, W.; Bhattacharya, D.; Davidson, B.L. Adenosine deamination in human transcripts generates novel microRNA binding sites. *Hum. Mol. Genet.* **2009**, *18*, 4801–4807. [[CrossRef](#)]
52. Available online: <https://www.ncbi.nlm.nih.gov/sra/?term=SRP098756> (accessed on 10 January 2021).
53. Available online: <http://picard.sourceforge.net/> (accessed on 15 March 2021).
54. Available online: <http://plants.ensem-bl.org/index.html> (accessed on 25 March 2021).
55. Available online: <http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi> (accessed on 13 May 2021).
56. Dai, X.; Zhuang, Z.; Zhao, P.X. psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **2018**, *46*, W49–W54. [[CrossRef](#)]
57. Available online: <https://swissmodel.expasy.org/> (accessed on 27 May 2021).