



Article

Integrated Analysis of Single-Molecule Real-Time Sequencing and Next-Generation Sequencing Reveals Insights into Drought Tolerance Mechanism of *Lolium multiflorum*

Qiuxu Liu [†], Fangyan Wang [†], Yang Shuai, Linkai Huang and Xinquan Zhang ^{*}

College of Grassland Science and Technology, Sichuan Agricultural University, Chengdu 611130, China; sicauliuqiuxu@163.com (Q.L.); wfy_zr@163.com (F.W.); syangerr@163.com (Y.S.); huanglinkai@sicau.edu.cn (L.H.)

* Correspondence: zhangxq@sicau.edu.cn

[†] These authors contributed equally to this work.

Abstract: *Lolium multiflorum* is widely planted in temperate and subtropical regions globally, and it has high economic value owing to its use as forage grass for a wide variety of livestock and poultry. However, drought seriously restricts its yield and quality. At present, owing to the lack of available genomic resources, many types of basic research cannot be conducted, which severely limits the in-depth functional analysis of genes in *L. multiflorum*. Therefore, we used single-molecule real-time (SMRT) and next-generation sequencing (NGS) to sequence the complex transcriptome of *L. multiflorum* under drought. We identified 41,141 DEGs in leaves, 35,559 DEGs in roots, respectively. Moreover, we identified 1243 alternative splicing events under drought. LmPIP5K9 produced two different transcripts with opposite expression patterns, possibly through the phospholipid signaling pathway or the negatively regulated sugar-mediated root growth response to drought stress, respectively. Additionally, 13,079 transcription factors in 90 families were obtained. An in-depth analysis of R2R3-MYB gene family members was performed to preliminarily demonstrate their functions by utilizing subcellular localization and overexpression in yeast. Our data make a significant contribution to the genetics of *L. multiflorum*, offering a current understanding of plant adaptation to drought stress.

Keywords: *Lolium multiflorum*; SMRT-Seq; drought stress; PacBio; R2R3-MYB



Citation: Liu, Q.; Wang, F.; Shuai, Y.; Huang, L.; Zhang, X. Integrated Analysis of Single-Molecule Real-Time Sequencing and Next-Generation Sequencing Reveals Insights into Drought Tolerance Mechanism of *Lolium multiflorum*. *Int. J. Mol. Sci.* **2022**, *23*, 7921. <https://doi.org/10.3390/ijms23147921>

Academic Editor: Sergey Shabala

Received: 8 June 2022

Accepted: 14 July 2022

Published: 18 July 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

Drought is a periodic and growing natural disaster that impacts extensive subject areas [1], including water resources [2,3] and crop yield [4–6], as well as a range of environmental systems [7], resulting in serious harm to ecological security and human society [8,9]. Over the past half-century, drought has become more and more serious all over the world [10], which has greatly reduced the productivity of grazing grasslands and artificial mowed grasslands.

L. multiflorum is one of the most important forage grasses and is widely grown in temperate and subtropical regions worldwide [11]. It is an excellent annual gramineous species recommended for planting, and its forage yield and quality cannot be replaced by other forage grasses at present [12]. It plays an important role in restoring degraded grasslands and establishing artificial grasslands and has both high ecological value and high economic benefit [12]. In recent years, its phytoremediation [13–15], bio-ethanol production [16] and anti-inflammatory medicinal properties [17] have also been reported. *L. multiflorum* is a highly self-incompatible plant, corresponding to the complex structure of its genome [18]. Meanwhile, the lack of available genomic resources hinders its improvement by breeders. Thus, there is an urgent need to construct valuable gene data sets and screen key candidate genes in *L. multiflorum*.

At present, vast quantities of data have been generated through second-generation high-throughput sequencing platforms. However, owing to the short reads generated by these technologies, it is quite difficult to obtain full-length sequences, and such sequencing technologies cannot work well in complex regions [19], which limits the ability of researchers to study gene function throughout the entire genome [20]. SMRT sequencing is a third-generation sequencing technology, which is more and more used in full-length sequencing [21,22]. It makes up for the deficiency from short reads and can generate full-length cDNA sequences (4–8 kb on average) without assembly or a reference genome, which greatly increases the potential for the discovery of genes and deeper studies of cell transcription [23,24]. There are many examples of SMRT sequencing being used to explore key genes or pathways to promote molecular breeding. In *Iris halophila*, metal ion transporters were found to be involved in the response to Pb stress using SMRT sequencing [25]. A key synthase in the benzyloisoquinoline alkaloid biosynthesis pathway, the main active substance in *Corydalis yanhusuo*, was identified by SMRT sequencing [26]. SMRT sequencing has become an ideal way to construct transcriptomes and analyze novel genetic material, especially for species with no reference genomes.

We utilized SMRT sequencing data generated with the PacBio Sequel platform, which produces long reads, to reveal full-length transcriptome information in *L. multiflorum* under drought. Such SMRT sequencing data can be complemented by NGS short reads [27]. Then, high-quality full-length transcript and drought-regulated genes in multiple tissues of *L. multiflorum* were obtained. On the basis of transcriptome data, we carried out functional annotation, lncRNA prediction, coding sequence prediction, TF analyses and functional validation of candidate genes by subcellular localization and overexpression in yeast. This is the first time to use SMRT sequencing to generate full-length transcription data from *L. multiflorum* under drought and is a very useful and important resource for further research of this important forage. This research can be utilized to elucidate the mechanisms of drought response of *L. multiflorum* and to create drought-tolerant, water-saving and environmentally friendly forage.

2. Results

2.1. Assembly of the Sequence Datasets and Functional Annotation

We obtained a total of 60.89 Gb of raw data from PacBio ISO-Seq and generated 13,787,867 subreads with a total size of 30.72 Gb. Furthermore, by self-correcting multiple single-molecule sequencing sequences, we obtained 1,026,983 CCS sequences after filtering. Based on their inclusion of 5'-primer, 3'-primer and poly-A tail sequences, CCS sequences were divided into full-length and non-full-length reads. A total of 879,040 full-length reads were found among the CCS sequences, and a total of 834,868 full-length non-chimera (FLNC) reads were obtained through ICE. Finally, LoRDEC software was used to identify sequencing errors, and the RNA-Seq short reads with high accuracy were combined for further correction. A total of 385,645 error-corrected consensus reads were obtained, with an average length of 2571 bp (Table 1, Figure 1C). Finally, CD-HIT software was used to eliminate redundancy among consensus reads, and 385,645 full-length non-redundant transcripts and 207,995 Unigenes of samples were obtained (Table 1). A total of 12,433 transcripts shared by the two samples were found by cluster analysis of transcripts with redundant sequences between the two samples (Figure 1A).

To further investigate gene function, non-redundant sequences were annotated by using CD-HIT. Overall, the transcripts with annotations corresponding to the Nr database were most common, with a total of 182,952 (Figure 1B). Additionally, 77,258 transcripts were annotated in all databases, and 189,898 transcripts were annotated in at least one database (Table S1).

Table 1. Summary of *Lolium multiflorum* single-molecule real-time sequencing results.

	CK		DR		Total
Subreads base (G)	15.48		15.24		30.72
number	6,944,546		6,843,321		13,787,867
Average length (bp)	2229		2227		2228
N50 (bp)	2652		2694		-
CCS	488,868		538,115		1,026,983
5'-primer	456,294		489,188		945,482
3'-primer	462,375		501,744		964,119
Poly-A	458,539		497,870		956,409
Full length	424,911		454,129		879,040
FLNC	403,543		431,325		834,868
	Before Correction	After Correction	Before Correction	After Correction	After Correction
Total nucleotides	469,861,260	471,334,422	519,052,371	520,093,777	991,428,199
Total number	184,267	184,267	201,378	201,378	385,645
Mean length (bp)	2550	2558	2578	2583	2571
Min length (bp)	192	193	200	197	193
Max length (bp)	14,449	14,437	14,422	14,348	14,437
N50 (bp)	2755	2759	2795	2798	-
N90 (bp)	1774	1778	1782	1784	-
Transcripts Length Interval	Number of Transcripts	Number of Unigenes	Number of Transcripts	Number of Unigenes	
<500 bp	1973	1086	418	211	
500–1k bp	4402	3289	3408	2299	
1–2k bp	46,268	17,602	50,843	25,263	
2–3k bp	80,106	40,389	87,946	45,276	
>3k bp	51,518	32,554	58,763	39,986	
Total	184,267	94,920	201,378	113,035	

To investigate whether drought stress affects AS in *L. multiflorum*, 26,169 non-redundant transcripts (Table S2) were processed into 10,629 UniTransModels (Table S2), and 1243 AS events (Table S2) were identified based on UniTransModels. Before and after drought stress, the occurrence of AS events was essentially uniform in both type and quantity. In descending order of quantity, the types can be ranked as follows: retained intron (RI), alternative 3' splice sites (A3), alternative 5' splice sites (A5), skipped exons (SE), alternative first exons (AF), and alternative last exons (AL).

2.2. Differentially Expressed Genes in *L. multiflorum* Leaves and Roots under Drought

To evaluate the reliability of the DEG analysis, Pearson correlation analysis was performed for pairs of samples. The coefficients of determination (R^2) among the three biological replicates in each condition were at least greater than 0.97, confirming the high correlation among biological replicates and the stability and reliability of the DEG analysis (Figure 2A). In leaves and roots of *L. multiflorum* under drought stress, 41,141 DEGs (19,155 down- and 21,986 up-regulated genes) (Figure 2C) and 35,559 DEGs (17,402 down- and 18,157 up-regulated genes) (Figure 2D) were found, respectively. All DEGs involved in the response to drought stress were analyzed, and 12 significant profiles were obtained by trend analysis (Figure 2B). Profile 15 represents 3955 genes up-regulated in response to drought stress in both leaves and roots (Figure S1A, Table S3). Profiles 14 and 11 represent 3993 up-regulated genes only in leaves and 3397 up-regulated genes only in roots, respectively (Figure S1B,C, Table S3). Profile 5 represents 4534 genes down-regulated in response to drought stress in both leaves and roots (Figure S1D, Table S3). These results indicated that there were more DEGs in *L. multiflorum* leaves under in vitro drought treatment and that the response of leaves to stress was broader and more active. At the same time, the trend analysis showed that more genes were down-regulated in both leaves and roots.

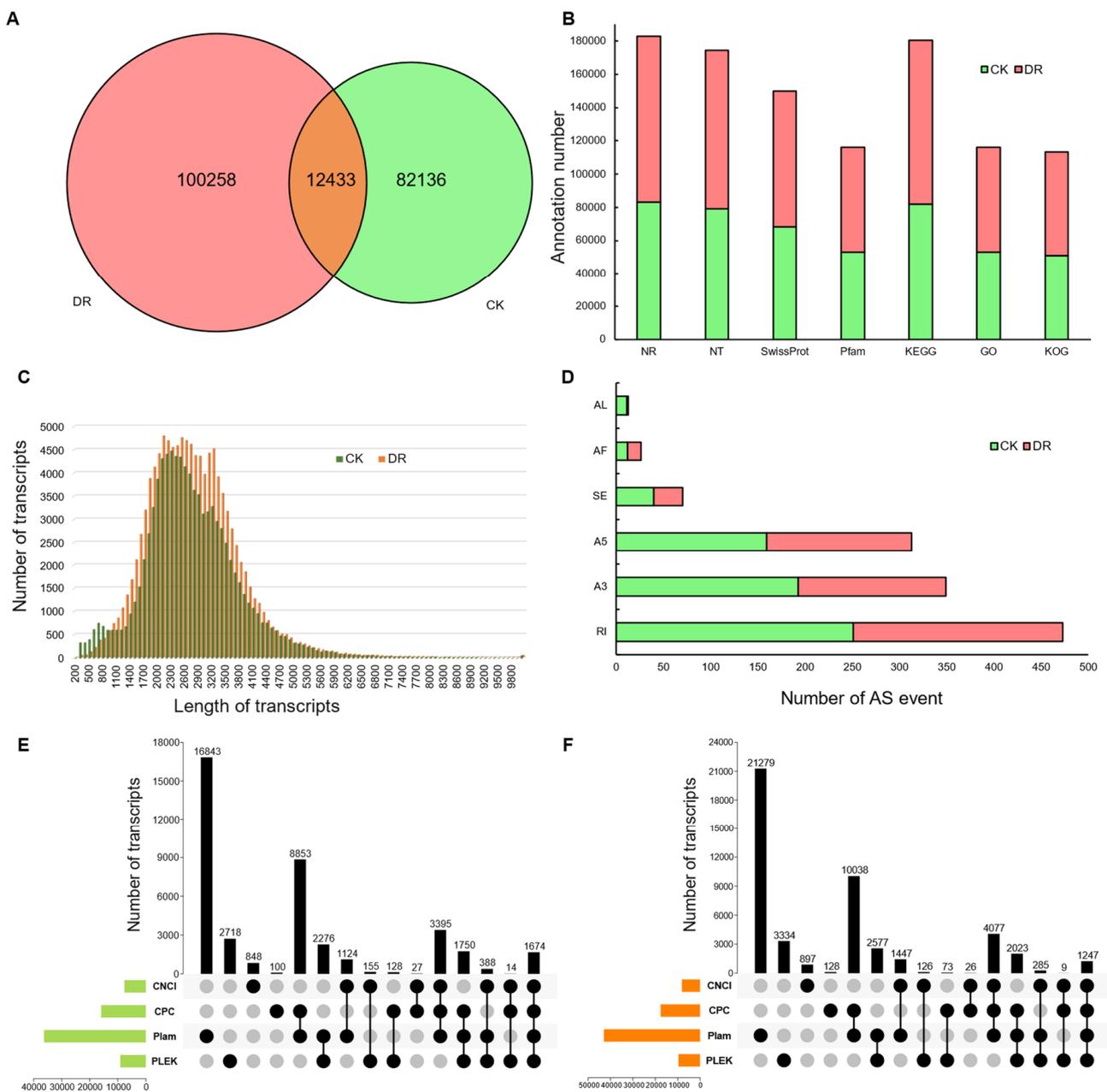


Figure 1. Gene function annotation and gene structure analysis. (A) Venn diagram comparing transcripts between control (CK) and drought (DR) conditions. (B) Summary of annotation results across the seven databases. (C) Length distribution of transcripts. (D) Statistical summary of alternative splicing (AS) events in the two samples. (E) UpSet graph showing lncRNA predictions for CK conditions. (F) UpSet graph showing lncRNA predictions for DR conditions.

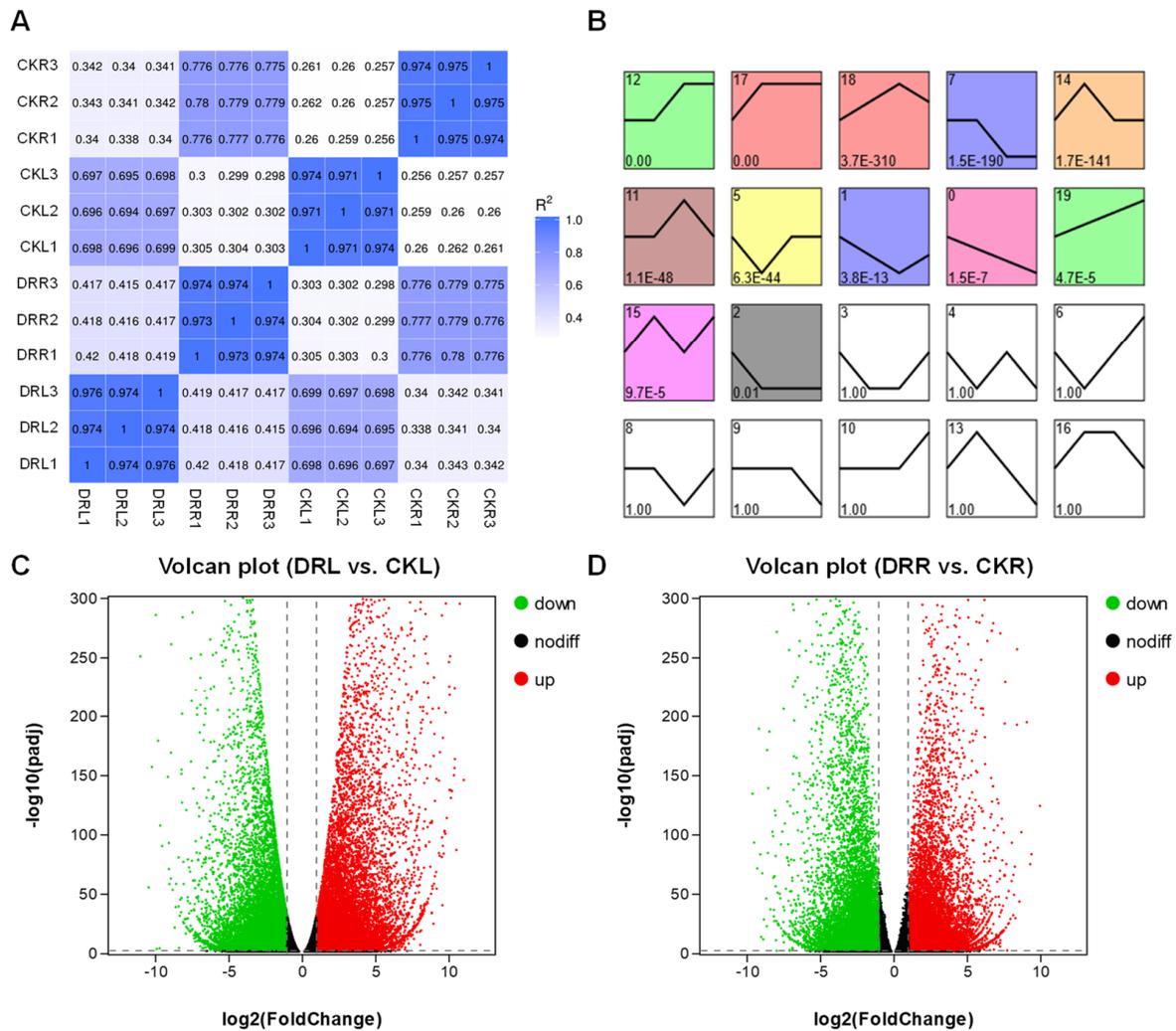


Figure 2. Differentially expressed genes in *L. multiflorum* leaves and roots under drought. **(A)** Pearson correlation coefficients between all samples. CKL and CKR represent leaves and roots under normal control conditions, respectively. DRL and DRR represent leaves and roots under drought stress, respectively. **(B)** Profiles ordered based on the P-value of genes assigned versus expected results. **(C)** Volcano plot of DRL versus CKL. **(D)** Volcano plot of DRR versus CKR. Green represents down-regulated expression, red represents up-regulated expression, and black indicates no significant difference.

A comparison of differences between the two treatment groups in leaves and roots (i.e., DRL versus CKL and DRR versus CKR) revealed that 11,940 DEGs responded to drought treatment in both tissues (Figure 3A, Table S4). To explore the biological functions of DEGs, GO and KEGG enrichment analyses were performed on all the differentially expressed genes in the two treatment groups (Figure 3B,C, Table S5). Among molecular function (MF) terms, ‘catalytic activity’ and ‘coenzyme binding’ were significantly enriched in both the DRL versus CKL and DRR versus CKR comparisons, simultaneously. Among cellular component (CC) terms, ‘membrane’, ‘1,3-beta-d-glucan synthase complex’, ‘integral component of membrane’, ‘intrinsic component of membrane’ and ‘membrane part’ were significantly enriched in the DRL versus CKL and DRR versus CKR comparisons, simultaneously. No biological process (BP) terms were enriched in both comparisons.

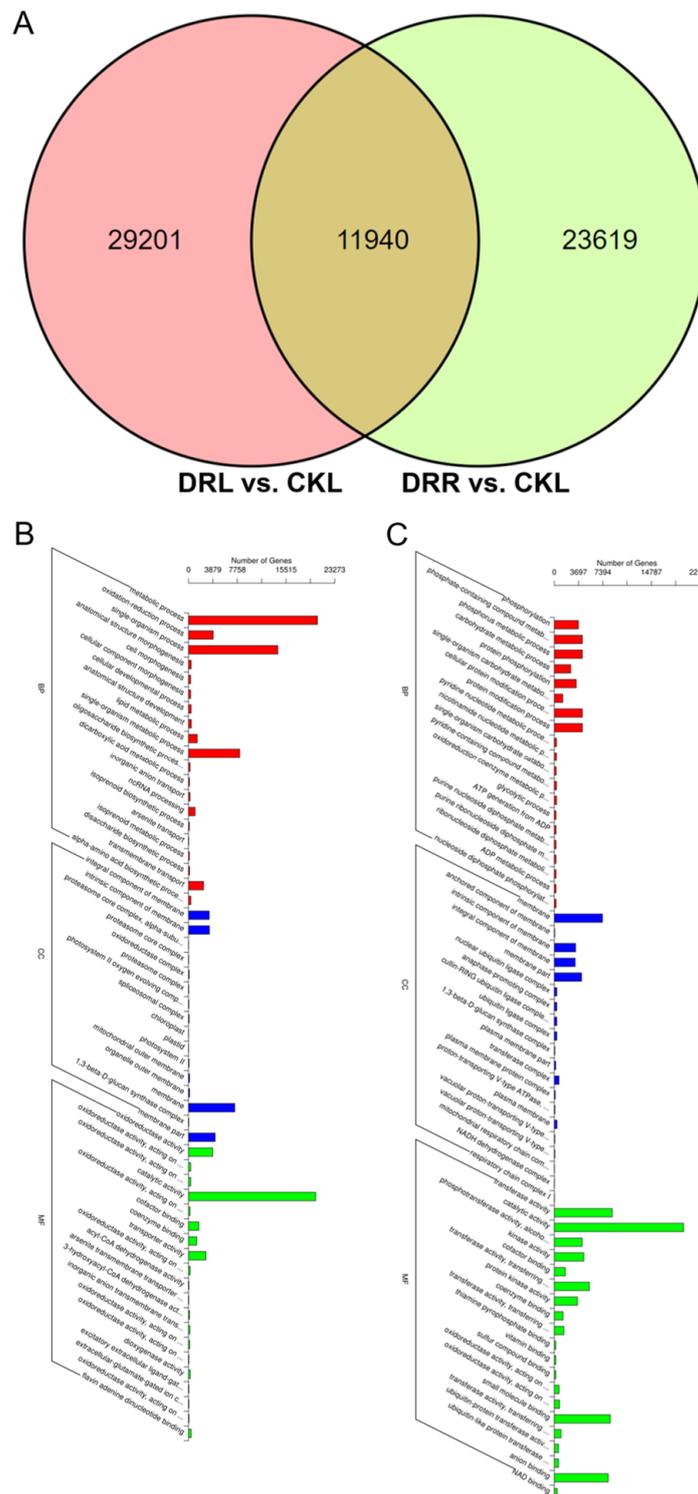


Figure 3. Biological function annotation of differentially expressed genes (DEGs). (A) Venn diagram for DEGs in the leaves in control versus drought conditions (DRL versus CKL) and roots in control versus drought conditions (DRR versus CKR) comparisons. (B) Gene Ontology (GO) enrichment analysis for DRL versus CKL. (C) GO enrichment analysis for DRR versus CKR.

2.3. Transcription Factor Statistics and Identification of R2R3-MYB Family Members

Transcription factors play an important role in regulating gene expression and have been an active research focus for decades. Full-length ryegrass transcriptome data were analyzed using the iTAK database, and the transcripts of 6512 transcription factors in

90 families were identified under CK conditions (Figure 4A, Table S6). A total of 6567 transcription factors in 88 families were identified under DR conditions (Figure 4B, Table S6). Under CK and DR conditions, the families represented by the most members were SNF2, SET and FAR1 and FAR1, C3H and SNF2, respectively. Under CK and DR conditions combined, four families ranked in the top ten in terms of the number of family members represented, namely SNF2, FAR1, MYB-related and C3H.

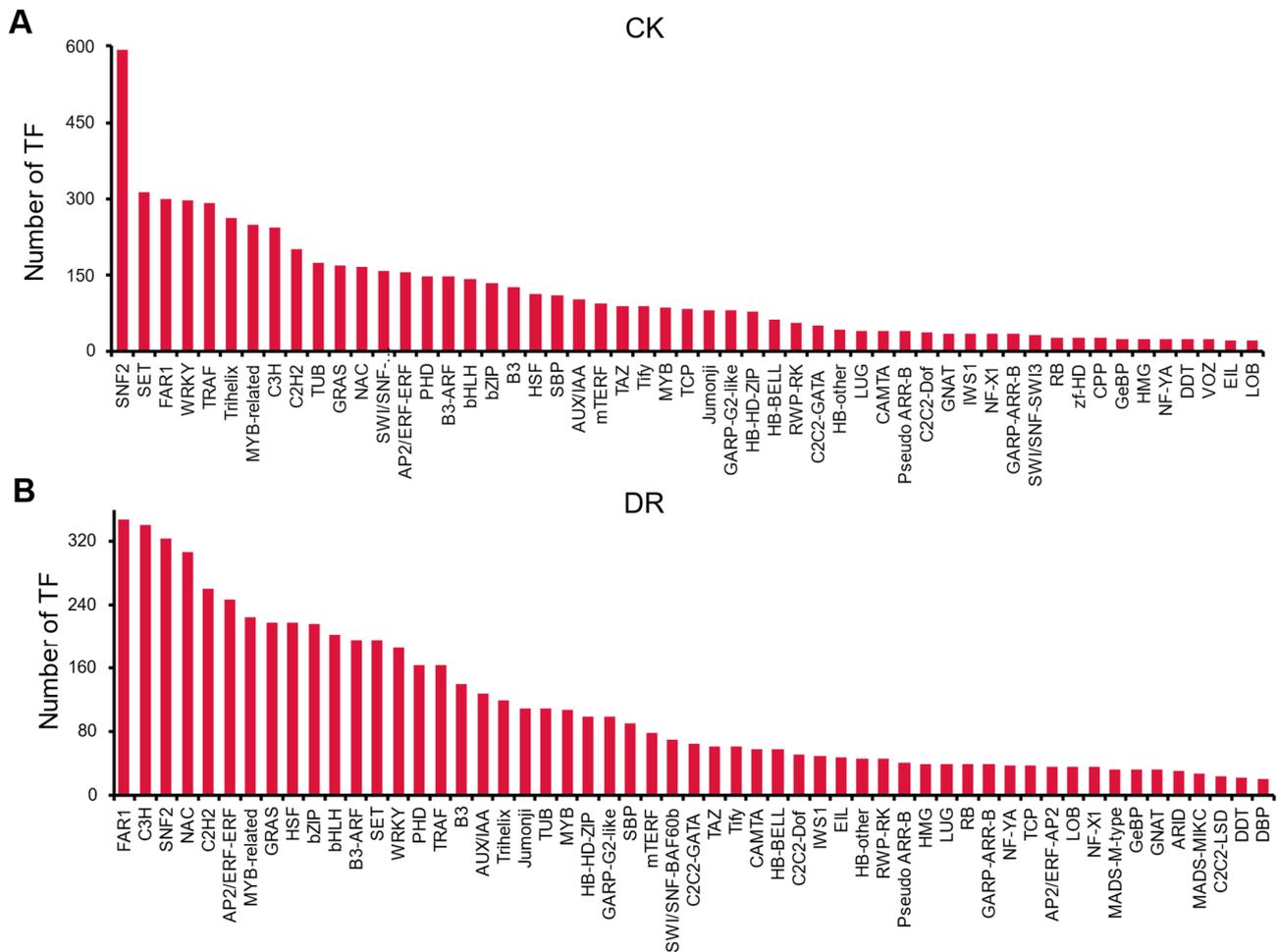


Figure 4. Summary of expressed transcription factor families. (A) Transcription factor families in CK. (B) Transcription factor families in DR. Transcription factor families are shown along the x-axis, while the corresponding numbers of different transcription factors are represented on the y-axis. Because of the large number of transcription factor families, the histogram shows only those families with more than 20 members.

Further identification of R2R3-MYB gene family members among MYB-related genes revealed 29 R2R3MYB genes expressed in *L. multiflorum*, sequentially named *LmMYB1* through *LmMYB29*. The ORF lengths of *LmMYB* members ranged from 720 to 2889 bp, and the predicted protein lengths ranged from 239 to 962 amino acids, with molecular weights ranging from 27.23 to 105.65 kDa, respectively (Table 2).

Table 2. R2R3-MYB family members identified in *Lolium multiflorum*.

	ORF Length (bp)	No. of AA	pI	Mw (kDa)
<i>LmMYB1</i>	747	248	9.12	27.75
<i>LmMYB2</i>	984	327	5.1	35.71
<i>LmMYB3</i>	1155	384	6.68	42.41
<i>LmMYB4</i>	1056	351	9.35	37.45
<i>LmMYB5</i>	849	282	5.08	31.61
<i>LmMYB6</i>	927	308	5.75	31.45
<i>LmMYB7</i>	906	301	5.26	33.52
<i>LmMYB8</i>	783	260	6.37	29.02
<i>LmMYB9</i>	1059	352	5.21	38.58
<i>LmMYB10</i>	720	239	7.13	27.23
<i>LmMYB11</i>	978	325	5.46	34.48
<i>LmMYB12</i>	1371	456	5.06	50.54
<i>LmMYB13</i>	1095	364	5.19	40.74
<i>LmMYB14</i>	1860	619	6.97	67.38
<i>LmMYB15</i>	1287	428	6.19	47.15
<i>LmMYB16</i>	873	290	6.39	32.21
<i>LmMYB17</i>	756	251	6.09	28.21
<i>LmMYB18</i>	915	304	5.27	33.87
<i>LmMYB19</i>	2094	697	4.83	75.73
<i>LmMYB20</i>	1872	623	6.97	67.75
<i>LmMYB21</i>	1089	362	5.18	40.47
<i>LmMYB22</i>	2889	962	5.11	105.65
<i>LmMYB23</i>	2538	845	5.4	93.78
<i>LmMYB24</i>	1002	333	5.21	37.37
<i>LmMYB25</i>	1650	549	5.15	59.42
<i>LmMYB26</i>	846	281	5.34	31.49
<i>LmMYB27</i>	1065	354	5.16	39.49
<i>LmMYB28</i>	2556	851	5.48	94.47
<i>LmMYB29</i>	933	310	5.11	34.54

Note: ORF, open reading frame; No. of AA, number of amino acids; Mw, molecular weight; pI, isoelectric point.

2.4. Genetic Analysis of Members of the R2R3-MYB Gene Family

To investigate the phylogenetic relationships of the R2R3MYB family members of *L. multiflorum* and *Arabidopsis thaliana*, a phylogenetic tree was constructed. Thus, R2R3MYB members from the two species could be divided into 13 subgroups, numbered S1 to S13, according to their phylogenetic relationships (Figure 5). In S1, there is only one gene family member, *LmMYB6*, and it is distantly related to all other members. Meanwhile, S3 and S5 subgroups only have members in *L. multiflorum*. All other subgroups consist of members in both *L. multiflorum* and *A. thaliana*.

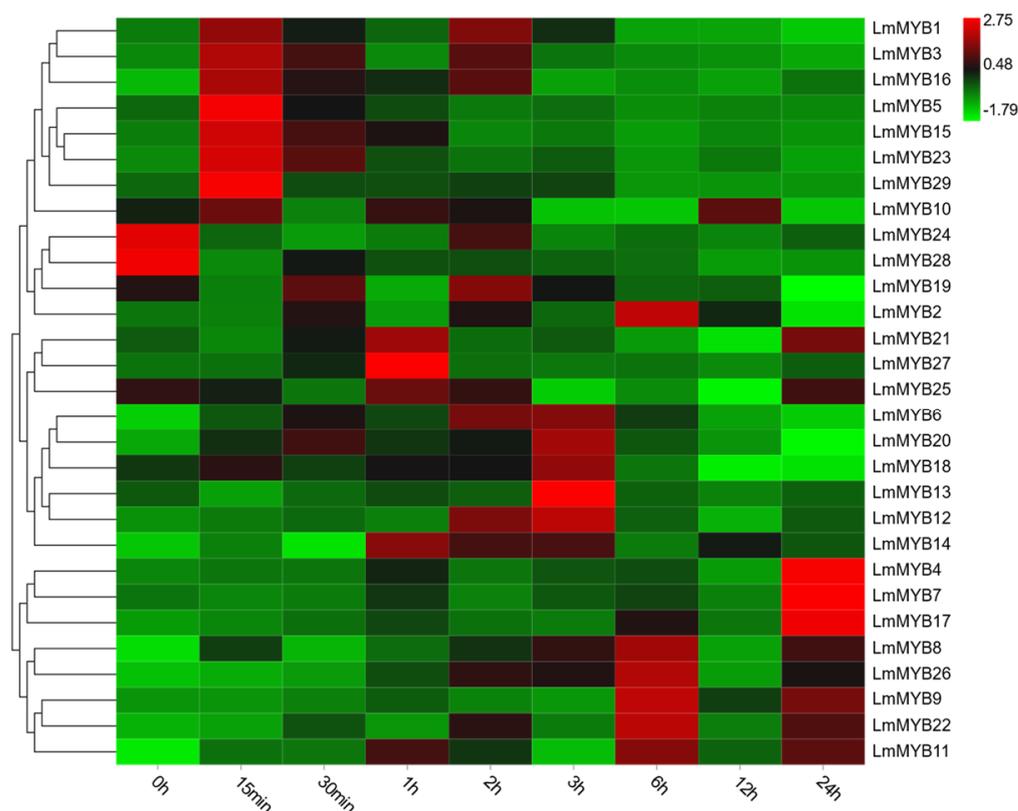


Figure 6. Expression profiles of *LmMYB* genes under drought stress conditions. Red and green represent relatively high and low expression compared to the control, respectively. The values used in the heat map were determined by qRT-PCR and normalized based on the expression of *eIF4A* and *HIS3*. Clustering was conducted by row.

2.6. *LmMYB* Transcripts Localized to the Nucleus Enhanced Abiotic Stress Tolerance in Yeast

The subcellular localization of *LmMYB1*, *LmMYB8* and *LmMYB9* in tobacco leaves was further studied, revealing that all three genes were expressed in the nucleus, which was consistent with the function of transcription factors (Figure 7A). To further characterize the response of *LmMYB1*, *LmMYB8* and *LmMYB9* to stress, we cloned them, inserted them into the *pYES2* vector, and then heterologously expressed them in the INVScI yeast line. All yeast-harboring empty vectors as well as factors containing *LmMYB1*, *LmMYB8* and *LmMYB9* were able to grow normally on SD-URA (2 g/L galactose) medium (Figure 7B). Only the INVScI strains with *LmMYB1*, *LmMYB8* and *LmMYB9* were able to grow on SD-URA (2 g/L galactose) medium under 3 M sorbitol (Figure 7C) and 1.5 M NaCl treatments (Figure 7D). This result indicated that overexpression of *LmMYB1*, *LmMYB8* and *LmMYB9* could significantly improve the resistance of INVScI yeast strains to osmotic stress compared with the strain harboring the empty *pYES2* vector. This also suggests that *LmMYB1*, *LmMYB8* and *LmMYB9* may be involved in the plant response to osmotic stress.

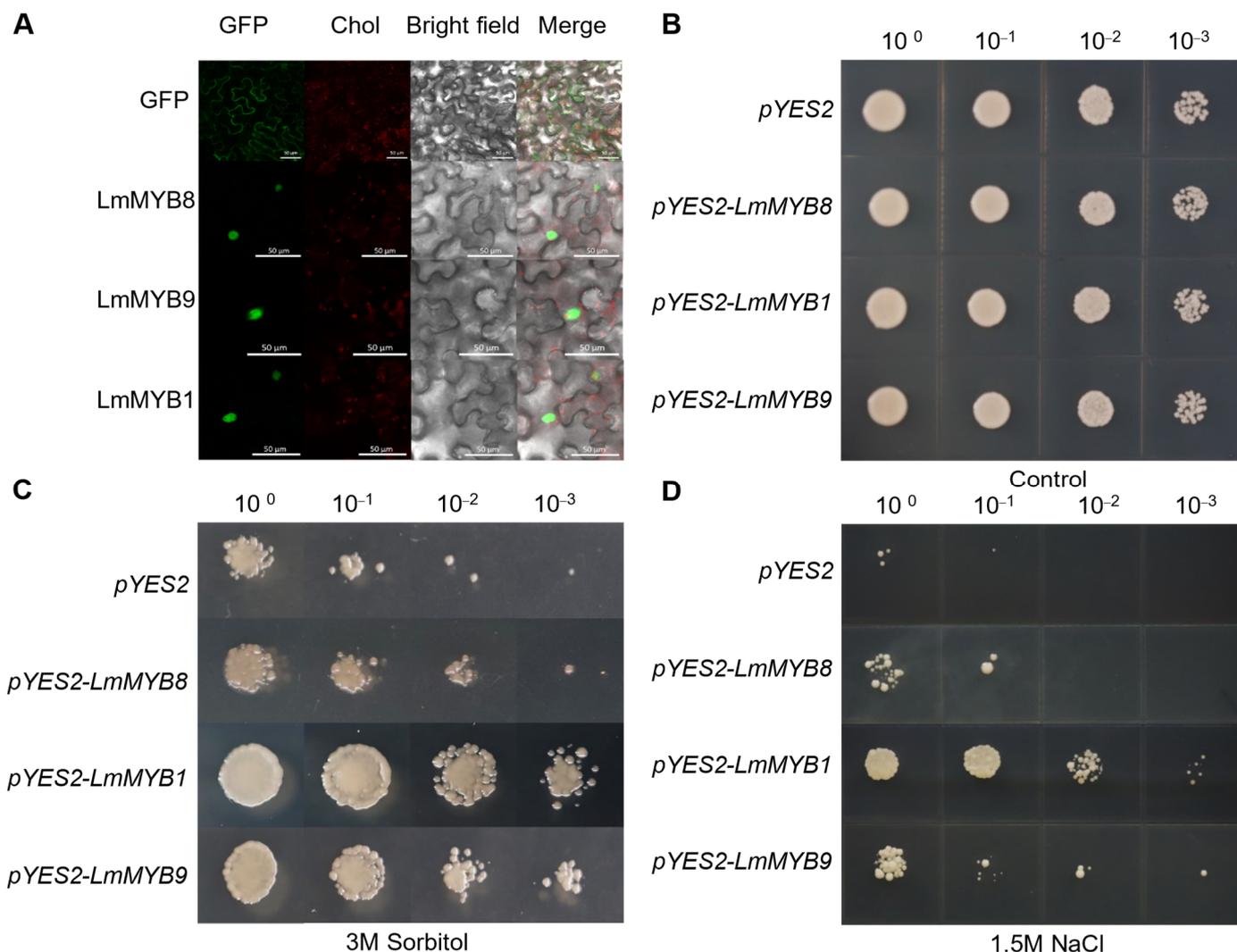


Figure 7. *LmMYB* transcripts are localized to the nucleus and enhanced drought and salt tolerance in yeast. (A) Subcellular localization of *LmMYB1*, *LmMYB8* and *LmMYB9*; scale bar, 50 μm. (B–D) The growth of yeast transformed with the empty vector *pYES2* or with *pYES* harbouring *LmMYB1*, *LmMYB8* or *LmMYB9* under control, 3 M sorbitol and 1.5 M NaCl conditions.

3. Discussion

As one of the most extensive forms of stress, drought has a huge impact on the survival of grasses. *L. multiflorum* is a vanguard grass species in artificial grassland construction, a preferred grass species for mowed grasslands and an important grass species for soil restoration. Understanding its response mechanism to drought stress is of substantial value for agricultural production, animal husbandry and ecological restoration. Currently, there is little available genomic information on *L. multiflorum*, so SMRT sequencing and NGS sequencing were used to construct a novel Unigene database for *L. multiflorum* under drought stress in this study. This extensive and comprehensive unigene database provides strong support for future research on the molecular mechanisms of drought responses in *L. multiflorum*. We identified 207,955 Unigenes and 1243 AS events and further characterized a number of key candidate genes involved in the response of *L. multiflorum* to drought stress, which can be used to advance the breeding of *L. multiflorum* to better adapt to increasingly arid environments.

3.1. A More Extensive and Complete Transcriptome Dataset

Compared to other *L. multiflorum* transcriptome studies using the Illumina platform, our results provide a more extensive and complete transcriptome dataset, with several critical strengths. First, our full-length transcriptome can later be used as a reference for annotating and assembling subsequent genomes of *L. multiflorum* and related species. Second, 207,955 accurate and high-quality full-length sequences were obtained, providing particularly valuable data for gene structure and gene function analyses. Third, the full-length transcripts generated in this study can be used to study the response of *L. multiflorum* to environmental changes with greater clarity. In this study, 207,955 Unigenes with an average length of 2571 bp were obtained by PacBio SMRT-Seq. Through the combination of Illumina and PacBio platform data, Unigenes are significantly increased in number and length compared with previous studies (Figure S2) [28,29]. The average predicted lengths of Unigenes using Illumina data were only 871 bp and 575.24 bp, respectively. The increase in number and length greatly enriched the *L. multiflorum* unigene library, making it both more extensive and more complete, thus providing solid data support for subsequent gene function studies.

3.2. Alternative Splicing Plays an Important Role in Complex Transcriptional Regulation

Alternative splicing plays an important role in stress responses and could enhance transcript diversity dramatically. The phospholipid signaling pathway is involved in regulating plant growth and aging [30]. Additionally, phosphatidylinositol phosphate 5-kinase (PIP5K) is a key part of the phospholipid signaling pathway [31]. PIP5K proteins play important roles in plants and have many functions. For example, *AtPIP5K1* is involved in responses to water stress and the abscisic acid signaling pathway in *A. thaliana* [32]. *AtPIP5K4* is involved in regulating stomatal opening in *A. thaliana* [33]. In this study, 1243 AS events were identified, and some notable phenomena were found. *LmPIP5K9* produced two different transcripts, i.e., CDS1 (i2_LQ_DE_c113853/f1p10/2988) and CDS2 (i3_LQ_DE_c41394/f1p1/3368), through alternative splicing of a retained intron (RI), each of which may perform different functions in *L. multiflorum* (Figure 8A). Read counts of the two transcripts indicated that the expression levels of CDS1 and CDS2 were low in both leaf and root samples under normal growth conditions, which was consistent with a recent study [34]. This finding implies that *LmPIP5K9* is not required under normal growth conditions. Additionally, CDS1 is rapidly upregulated, about four-fold, after induction by drought stress, which is also corroborated by other studies [34,35]. However, the expression of CDS2, which itself is low under normal growth conditions, is further down-regulated and almost not expressed at all under drought stress. Thus, two transcripts of the same gene have different expression patterns. Accordingly, *LmPIP5K9* may up-regulate the CDS1-mediated phospholipid signaling pathway to participate in the drought stress response. In contrast, CDS2 interacts with a cytosolic invertase to negatively regulate sugar-mediated root growth, as previously reported [36]. In future work, we will confirm the functions of these two transcripts, initially in *A. thaliana*.

3.3. Key Distinctive Candidate Genes Involved in the *L. multiflorum* Drought Stress Response

Drought stress greatly affected gene expression in *L. multiflorum*, and the affected genes exhibited different expression patterns. We combined differential gene analysis, alternative splicing analysis and gene family analysis to identify some distinctive candidate genes involved in the response to drought stress. Homeobox (HOX) genes have been identified and characterized in many eukaryotes and are involved in regulating various aspects of growth and development [37]. We found that HOX22 (i1_LQ_DE_c57408/f1p2/1087) and HOX24 (i3_HQ_DE_c33357/f3p2/3113) had higher expression levels in DRL samples than other samples (Figure 8B). At the same time, there was an enrichment of DEGs in the Intrinsic/Integral component of Golgi membrane identified by GO analysis, suggesting that HOX22/24 responds to drought stress by participating in the formation of the Golgi apparatus. High-affinity K⁺ transporters (HAK) are present in all plants with known

genomes, but not in animals [38]. These proteins play an important role in K⁺ uptake and transport [39,40] and are also involved in osmotic regulation [41], stabilizing plants by balancing K⁺ homeostasis during cell growth and drought stress responses. In the present study, the expression of *LmHAK7* in root tissues after drought stress was remarkably high compared with leaf tissues (Figure 8C), which was consistent with a recent study [42]. Characterizing intense responses to drought stress (i.e., fold change DRR:CKR = 71.5) and ultra-high expression in roots (i.e., fold change DRR:DRL = 9.45) will be important focuses of our future work. In the present study, we also found that *LmMYB1*, *LmMYB8* and *LmMYB9* were able to improve drought tolerance and salt tolerance in yeast (Figure 7), suggesting that they might have the same function in plants. Future experiments will aim to genetically transform *A. thaliana* and *L. multiflorum* for functional verification.

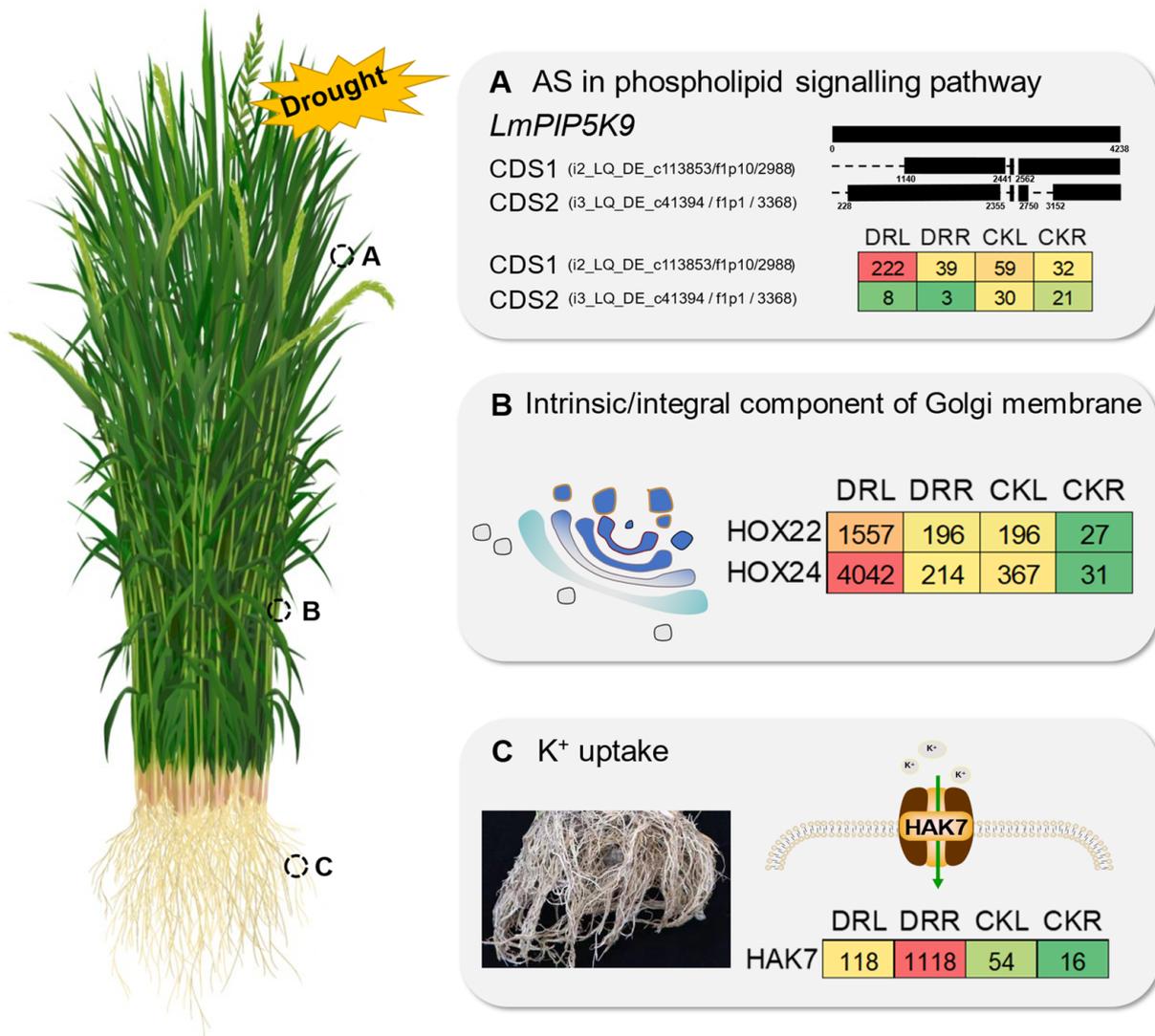


Figure 8. The complexity of transcriptional regulation in the *Lolium multiflorum* drought response. (A) The role of alternative splicing (AS) in the phospholipid signaling pathway. The black bars represent sequences, and the grey dotted lines represent gaps. (B) The HOX gene family is highly expressed in leaf tissues under drought and participates in intrinsic/integral components of the Golgi membrane. (C) The plant-specific gene HAK was highly expressed in root tissue under drought and is involved in K⁺ uptake. Red and green represent high and low expression, respectively, and the number in each box represents the corresponding read count.

4. Materials and Methods

4.1. Material Cultivation and Sample Collection

In this study, *L. multiflorum* cultivar 'Chuannong No. 1' was used. The experimental subjects for transcriptional sequencing included 300 individual plants during their flowering period, all of which were cultivated at the Ya'an research station of Sichuan Agricultural University. In order to obtain more comprehensive full-length transcriptome sequence information, roots, stems, leaves and flowers were collected from 30 randomly selected individual plants in the control group (CK) and in vitro drought treatment group after 6 h (DR).

Samples for expression pattern analysis were planted in pots (25 × 19 × 6 cm) in a growth chamber with day/night cycles of 16 h/8 h and 22/20 °C. When the number of leaves on plants reached three or four, stress treatments were initiated. The 15% PEG6000 is used to simulate drought stress. The leaves were sampled at 0 h, 15 min, 30 min, 1 h, 2 h, 3 h, 6 h, 12 h and 24 h after drought stress. All tissues were immediately stored in liquid nitrogen after sampling and then stored at −80 °C.

4.2. Illumina cDNA Library Construction and Next-Generation Sequencing

After RNA was extracted from all samples, 3 µg of RNA was used for Illumina cDNA library construction. Total RNA was extracted by grinding tissue in TRIzol reagent (Invitrogen, Carlsbad, CA, USA) on dry ice and processed following the protocol provided by the manufacturer. For NGS, a total of 12 Illumina cDNA libraries were constructed by using the Illumina Stranded RNA Library Prep Kit (New York, NY, USA). Six libraries were constructed from root tissue and the other six were from leaf tissue. The cDNA library samples were sequenced by Gene Denovo Biotechnology Co. (Guangzhou, China). The quality control of the NGS raw data was conducted as follows: the sequenced raw reads with the adapter and primer sequences and poly-N were filtered out. The clean data were used to correct the PacBio sequencing data in the next step.

4.3. PacBio cDNA Library Construction and Single-Molecule Real-Time (SMRT) Sequencing

After RNA was extracted from all samples, 1 µg of RNA was used for PacBio cDNA library construction. Total RNA was extracted from each sample by using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). We constructed two libraries, one from a mixture of roots, stems, leaves and flowers under drought treatment, and the other from a mixture of roots, stems, leaves and flowers under the control treatment. The SMRT sequencing library was prepared with the Clontech SMARTer PCR cDNA Synthesis Kit and the BluePippin Size Selection System protocol. The mRNA enriched by Oligo (dT) magnetic beads was reverse transcribed into cDNA. Then, double-stranded cDNA was generated with the optimum cycle number. In addition, >4 kb size selection was performed using the BluePippin™ Size-Selection System. Then, large-scale PCR was performed for the subsequent SMRTbell library construction. The SMRTbell library was sequenced by Gene Denovo Biotechnology Co. (Guangzhou, China).

The PacBio Sequel sequencing platform, based on SMRT-Seq technology, was used to generate full-length sequence data. The whole data processing process is carried out based on SMRTlink 5.1 (<https://www.pacb.com> (accessed on 20 May 2020)). Circular consensus sequence (CCS) data were generated from the subread BAM files. CCS.BAM files that were output were then classified into full-length and non-full-length reads using pbclassify.py. The non-full length and full-length fasta files produced were then fed into the cluster step, which performs isoform-level clustering (ICE), followed by final Arrow polishing. Additional nucleotide errors in consensus reads were corrected using Illumina RNA-Seq data with the software LoRDEC. Any redundancy in corrected consensus reads was removed by CD-HIT to obtain final transcripts for the subsequent analysis (Figure 9).

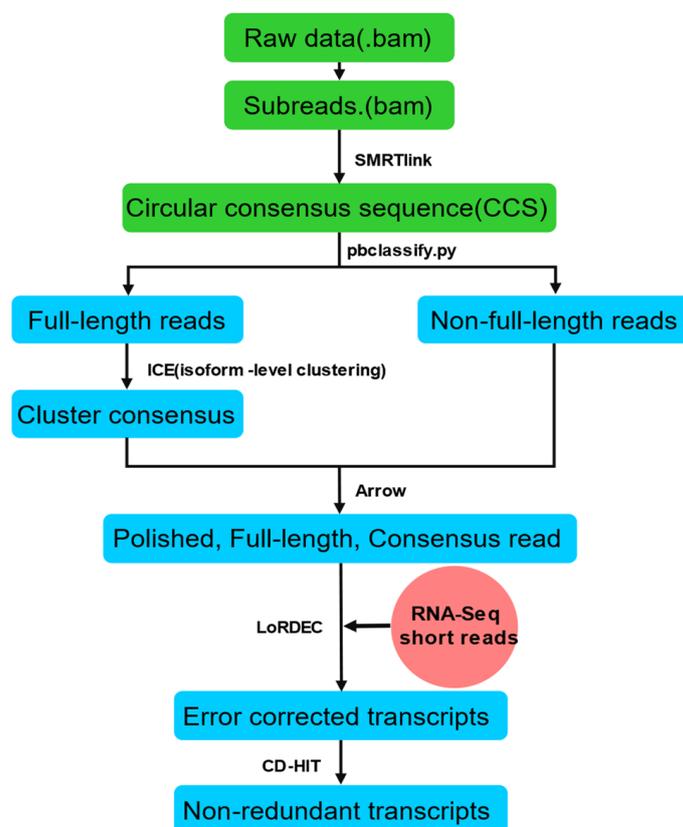


Figure 9. Analysis pipeline for SMRT-Seq data.

4.4. Functional Annotation of PacBio Isoforms

To annotate the isoforms identified in the sequencing data, isoforms were BLASTed against the NCBI non-redundant protein (Nr) database (<http://www.ncbi.nlm.nih.gov> (accessed on 20 May 2020)), NCBI nucleotide (Nt) database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://www.genome.jp/kegg> (accessed on 20 May 2020)), the Swiss-Prot protein database (<http://www.expasy.ch/sprot> (accessed on 20 May 2020)) and the COG/KOG database (<http://www.ncbi.nlm.nih.gov/COG> (accessed on 20 May 2020)) with the BLASTx program (<http://www.ncbi.nlm.nih.gov/BLAST/> (accessed on 20 May 2020)) at an E-value threshold of 10^{-5} to evaluate sequence similarity with genes of other species. GO annotation was analyzed using Blast2GO software [43] with Nr annotation isoform results. Then, functional classification of isoforms was performed using WEGO software [44].

4.5. Identification of Alternative Splicing Events and lncRNA Prediction

The non-redundant transcripts were processed with the Coding GENome reconstruction Tool [45]. Each transcript family was further reconstructed into one or more unique transcript model(s) (referred to as UniTransModels). Error-corrected non-redundant transcripts were mapped to UniTransModels. Splicing junctions for transcripts mapped to the same UniTransModels were examined, and transcripts with the same splicing junctions were collapsed. Collapsed transcripts with different splicing junctions were detected as transcription isoforms of UniTransModels. Alternative splicing (AS) events were identified with SUPPA (<https://github.com/comprna/SUPPA> (accessed on 20 May 2020)) [46].

We used the Coding-Non-Coding-Index (CNCI) [47], Coding Potential Calculator (CPC) [48], Pfam-scan [49] and PLEK [50] software tools to predict lncRNA. CNCI profiles adjoin nucleotide triplets to effectively distinguish protein-coding from non-coding sequences independent of known annotations. CPC mainly assesses the extent and quality of the ORF in a transcript and searches the sequences in known protein sequence databases

to clarify the coding and non-coding transcripts. Each transcript was translated in all three possible frames and mapped in the Pfam database using Pfam Scan. Any transcript with a Pfam hit was excluded in the following steps. The PLEK support vector machine classifier uses an optimized K-mer approach to construct the best classifier to assess the coding potential for species that lack high-quality genome sequences and annotations. Default parameters were used for four tools. Transcripts predicted with protein-coding potential by any or all of the three tools above were filtered out, and those without any identified coding potential were our candidate set of lncRNAs.

4.6. Quantification of Gene Expression Levels and Differential Expression Analysis

Reference sequences were mapped prior to the quantification of gene expression. The reference sequence mapping used CD-HIT to identify the transcripts that are redundant after generating the corrected consensus sequence. The clean reads from RNA-Seq were mapped to the reference sequence. The process used RSEM software [51], and the parameters for bowtie2 were end-to-end and sensitive mode. Default parameters were used for all other settings.

The read count for each transcript was obtained from the mapping results. Differential expression analysis of the two groups was performed using the DESeq2 R package (1.34.0) (<http://bioconductor.org/packages/release/bioc/html/DESeq2.html> (accessed on 20 May 2020)). The resulting P-values were adjusted using Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted *p*-value < 0.05 as determined by DESeq were identified as differentially expressed.

GO enrichment analysis of differentially expressed genes (DEGs) was conducted using the Goseq R package (1.46.0) (<http://bioconductor.org/packages/release/bioc/html/goseq.html> (accessed on 20 May 2020)). GO terms with corrected *p*-values < 0.05 were considered significantly enriched among DEGs. KEGG enrichment analysis was conducted as implemented in KOBAS 3.0 (<http://kobas.cbi.pku.edu.cn/kobas3> (accessed on 20 May 2020)) with its default parameters.

4.7. Transcription Factor Analysis and Identification of R2R3-MYB Gene Family Members

Transcription factors (TFs) are groups of protein molecules that can bind to specific sequences in the 5' upstream sequences of genes to ensure that the target genes are expressed at a specific intensity at a specific time and location [52]. We used iTAK software [53] to predict plant TFs. R2R3-MYB genes were identified in the PFAM protein family database using HMMER 3.0 software [54], with the MYB-like DNA-binding domains (Pfam, PF00249) as the search query [55] and an initial threshold value of $E \leq 10^{-10}$. Basic information about these genes, including PIs, MWs and subcellular localization, was predicted using the ExpASy website (<https://web.expasy.org/protparam/> (accessed on 10/02/2021)) [56].

4.8. Phylogenetic Analysis

The inferred protein sequences of the R2R3-MYB from *L. multiflorum* and *A. thaliana* were aligned using ClustalW with its default parameters. MUSCLE [57] and Clustal Omega [58] were also used to verify alignment results. The proteins of *A. thaliana* MYB were downloaded from the Arabidopsis Information Resource (TAIR) (<http://www.aabidopsis.org/> (accessed on 10 February 2021)). Based on the alignment of the MYB domain, a phylogenetic tree was constructed with MEGA 7.0 [59] using the neighbor-joining (NJ) method. Bootstrap values (>50%) were estimated using 1000 replicates. Interactive Tree Of Life (iTOL) software was used to optimize the obtained phylogenetic tree [60].

4.9. Expression Pattern Analysis of the R2R3-MYB Gene Family

Total RNA was extracted from leaves under drought stress. The cDNA was synthesized using a MonScript™ RTIII Super Mix with dsDNase (Two-Step) Kit from Monad Biotech Co., Ltd. (Suzhou, China). Subsequently, real-time quantitative PCR (qRT-PCR) was performed using SYBR qPCR Master Mix (Vazyme, China). Gene-specific primers

(shown in Supplementary Table S1) were designed to avoid the conserved region. Both eIF4A and HIS3 were used as reference genes [61].

4.10. Subcellular Localization and Heterologous Expression

Subcellular localization prediction of gene family members was conducted using the ExPASy website. To confirm the predicted subcellular localization, we inserted the full-length sequences of *LmMYB1*, *LmMYB8* and *LmMYB9* into the pYBA1132 vector. Then, all of them were transformed into tobacco leaves by Agrobacterium-mediated transformation. The primers used are listed in Supplementary Table S7. Empty vector GFP was used as a control. The fluorescence images of GFP fusion proteins were observed by FV10 confocal microscopy.

4.11. Heterologous Expression of *LmMYB1*, *LmMYB8* and *LmMYB9* in Yeast

LmMYB1, *LmMYB8* and *LmMYB9* were amplified from the PacBio cDNA library of *L. multiflorum* constructed using the primers listed in Supplementary Table S7. The correct CDS regions were cloned into a pYES2 vector for expression in yeast. Cells of the sorbitol- and NaCl-sensitive yeast strain INVSc1 were obtained from MiaoLing Plasmid Platform. The plasmids empty pYES2, pYES2-*LmMYB1*, pYES2-*LmMYB8* and pYES2-*LmMYB9* were transformed into INVSc1 yeast by Carrier DNA (Vazyme, China).

Yeast transformants were screened using SD-Ura plates. The whole process lasted for 2 days at 28 °C. To analyze sorbitol and NaCl resistance, the validated single colonies were cultured in liquid SD-Ura medium (2% galactose) at 28 °C and 150 rpm. When the medium OD600 value reaches 2.0, gradient dilution is performed (10^{-1} , 10^{-2} , 10^{-3}). Diluted yeast was spotted onto SD-Ura plates containing 3 M sorbitol or 2 M NaCl in turn, which were photographed and observed by the naked eye.

5. Conclusions

Through the analysis of transcriptome changes in leaves and roots of annual ryegrass under drought stress using SMRT and NGS sequencing technologies we identified many genes with potentially related functions and revealed complex transcript responses such as alternative splicing and lncRNA expression. These results contribute to the current understanding of the complexity of transcriptional regulation in plant drought responses. In particular, the R2R3-MYB transcription factor family was identified and analyzed using the high-quality full-length transcriptome. The candidate drought response regulatory factors *LmMYB1*, *LmMYB8* and *LmMYB9* in *L. multiflorum* were preliminarily verified in both tobacco and yeast. These results help clarify the mechanism of plant responses to stress. Future studies will focus on plant responses to stress to improve our understanding of plant adaptation to climate change.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23147921/s1>.

Author Contributions: X.Z. and L.H. contributed to the conception of the study; Q.L., F.W. and Y.S. performed the experiment; Q.L. and F.W. contributed significantly to analysis and manuscript preparation; Q.L. and F.W. performed the data analyses and wrote the manuscript; Q.L., X.Z. and L.H. helped perform the analysis with constructive discussions. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by the China Agriculture Research System of the Ministry of Finance (MOF) and the Ministry of Agriculture and Rural Affairs (MARA) (CARS-34), a Sichuan Province Breeding Research grant, and the National Project on Sci-Tec Foundation Resources Survey (2017FY100602).

Data Availability Statement: The datasets generated in this study have been uploaded to the NCBI database under the accession number PRJNA634598.

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

References

1. Vicente-Serrano, S.M. Foreword: Drought complexity and assessment under climate change conditions. *Cuad. Investig. Geográfica* **2016**, *42*, 7–11. [[CrossRef](#)]
2. Folland, C.K.; Hannaford, J.; Bloomfield, J.P.; Kendon, M.; Wallace, E. Hydrology and Earth System Sciences. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 2353–2375. [[CrossRef](#)]
3. Lanen, H.V.; Wanders, N.; Tallaksen, L.M.; Loon, A.V. Hydrological drought across the world: Impact of climate and physical catchment structure. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1715–1732. [[CrossRef](#)]
4. Kim, W.; Iizumi, T.; Nishimori, M. Global Patterns of Crop Production Losses Associated with Droughts from 1983 to 2009. *J. Appl. Meteorol. Clim.* **2019**, *58*, 1233–1244. [[CrossRef](#)]
5. Pea-Gallardo, M.; Vicente-Serrano, S.M.; Hannaford, J.; Lorenzo-Lacruz, J.; Kenawy, A.E. Complex influences of meteorological drought time-scales on hydrological droughts in natural basins of the contiguous United States. *J. Hydrol.* **2018**, *568*, 611–625. [[CrossRef](#)]
6. Webber, H.; Ewert, F.; Olesen, J.E.; Müller, C.; Fronzek, S.; Ruane, A.C.; Bourgault, M.; Martre, P.; Ababaei, B.; Bindi, M. Diverging importance of drought stress for maize and winter wheat in Europe. *Nat. Commun.* **2018**, *9*, 4249. [[CrossRef](#)]
7. Vicente-Serrano, S.M.; Quiring, S.M.; Pea-Gallardo, M.; Yuan, S.; Domínguez-Castro, F. A review of environmental droughts: Increased risk under global warming? *Earth Sci. Rev.* **2019**, *201*, 102953. [[CrossRef](#)]
8. Törnros, T.; Menzel, L. Addressing drought conditions under current and future climates in the Jordan River region. *Hydrol. Earth Syst. Sci.* **2013**, *18*, 305–318. [[CrossRef](#)]
9. Giddens, A. *The Politics of Climate Change*, 2nd ed.; Polity: Cambridge, UK, 2011.
10. Doede, A.L.; Deguzman, P.B. The Disappearing Lake: An Historical Analysis of Drought and the Salton Sea in the Context of the GeoHealth Framework. *GeoHealth* **2020**, *4*, e2020GH000271. [[CrossRef](#)]
11. Knorst, V.; Yates, S.; Byrne, S.; Asp, T.; Widmer, F.; Studer, B.; Kölliker, R. Liker First assembly of the gene-space of *Lolium multiflorum* and comparison to other Poaceae genomes. *Grassl. Sci.* **2019**, *65*, 125–134. [[CrossRef](#)]
12. Xin-Yue, Z.; Yuan-Hua, L.L.; Wen-Long, G.; Rui-Zhen, Z. Research development of Italian ryegrass. *Pratacultural Sci.* **2009**, *26*, 55–60.
13. Fang, Z.; Hu, Z.; Zhao, H.; Yang, L.; Ding, C.; Lou, L.; Cai, Q. Screening for cadmium tolerance of 21 cultivars from Italian ryegrass (*Lolium multiflorum* Lam) during germination. *Grassl. Sci.* **2017**, *63*, 36–45. [[CrossRef](#)]
14. Mugica-Alvarez, V.; Cortés-Jiménez, V.; Vaca-Mier, M.; Domínguez-Soria, V. Phytoremediation of Mine Tailings Using *Lolium Multiflorum*. *Int. J. Environ. Sci. Dev.* **2015**, *6*, 246–251. [[CrossRef](#)]
15. Liu, Z.; He, X.; Chen, W.; Zhao, M. Ecotoxicological responses of three ornamental herb species to cadmium. *Environ. Toxicol. Chem.* **2013**, *32*, 1746–1751. [[CrossRef](#)] [[PubMed](#)]
16. Yasuda, M.; Takenouchi, Y.; Nitta, Y.; Ishii, Y.; Ohta, K. Italian ryegrass (*Lolium multiflorum* Lam) as a High-Potential Bio-Ethanol Resource. *Bioenerg. Res.* **2015**, *8*, 1303–1309. [[CrossRef](#)]
17. Choi, K.C.; Son, Y.O.; Hwang, J.M.; Kim, B.T.; Lee, J.C. Antioxidant, anti-inflammatory and anti-septic potential of phenolic acids and flavonoid fractions isolated from *Lolium multiflorum*. *Pharm. Biol.* **2017**, *55*, 611–619. [[CrossRef](#)]
18. Cornish, M.A.; Hayward, M.D.; Lawrence, M.J. Self-incompatibility in ryegrass. *Heredity* **1979**, *43*, 95–106. [[CrossRef](#)]
19. Morganti, S.; Tarantino, P.; Ferraro, E.; D’Amico, P.; Viale, G.; Trapani, D.; Duso, B.A.; Curigliano, G. Complexity of genome sequencing and reporting: Next generation sequencing (NGS) technologies and implementation of precision medicine in real life. *Crit. Rev. Oncol. Hematology* **2019**, *133*, 171–182. [[CrossRef](#)]
20. Au, K.F.; Sebastiano, V.; Afshar, P.T.; Durruthy, J.D.; Lee, L.; Williams, B.A.; Van Bakel, H.; Schadt, E.E.; Reijo-Pera, R.A.; Underwood, J.G. Characterization of the human ESC transcriptome by hybrid sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, e4821–e4830. [[CrossRef](#)]
21. Chaisson, M.J.P.; Huddleston, J.; Megan, Y.D.; Sudmant, P.H.; Malig, M. Resolving the complexity of the human genome using single-molecule sequencing. *Nature* **2015**, *517*, 608–611. [[CrossRef](#)]
22. Chen, X.; Bracht, J.; Goldman, A.; Dolzhenko, E.; Clay, D.; Swart, E.; Perlman, D.; Doak, T.; Stuart, A.; Amemiya, C. The Architecture of a Scrambled Genome Reveals Massive Levels of Genomic Rearrangement during Development. *Cell* **2014**, *158*, 1187–1198. [[CrossRef](#)] [[PubMed](#)]
23. Ju, C.; Zhao, Z.; Wei, W. Efficient approach to correct read alignment for pseudogene abundance estimates. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2016**, *14*, 522–533. [[CrossRef](#)] [[PubMed](#)]
24. Sharon, D.; Tilgner, H.; Grubert, F.; Snyder, M. A single-molecule long-read survey of the human transcriptome. *Nat. Biotechnol.* **2013**, *31*, 1009–1014. [[CrossRef](#)] [[PubMed](#)]
25. Qla, B.; Yza, B.; Ywa, B.; Cga, B.; Sha, B.; Opd, C.; Hya, B. Combining single-molecule sequencing and next-generation sequencing to provide insight into the complex response of *Iris halophila* Pall. to Pb exposure. *Ind. Crop. Prod.* **2021**, *168*, 113623.
26. Xu, D.; Lin, H.; Tang, Y.; Huang, L.; Xu, J.; Nian, S.; Zhao, Y. Integration of full-length transcriptomics and targeted metabolomics to identify benzyloquinoline alkaloid biosynthetic genes in. *Hortic. Res.* **2021**, *8*, 16. [[CrossRef](#)]
27. 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](#)]

28. Cechin, J.; Piasecki, C.; Benemann, D.P.; Kremer, F.S.; Galli, V.; Maia, L.C.; Agostinetto, D.; Vargas, A.L. *Lolium multiflorum*-Transcriptome Analysis Identifies Candidate Target Genes Involved in Glyphosate-Resistance Mechanism in. *Plants* **2020**, *9*, 685. [[CrossRef](#)]
29. Pan, L.; Zhang, X.; Wang, J.; Ma, X.; Zhou, M.; Huang, L.; Nie, G.; Wang, P.; Yang, Z.; Li, J. Transcriptional Profiles of Drought-Related Genes in Modulating Metabolic Processes and Antioxidant Defenses in *Lolium multiflorum*. *Front. Plant Sci.* **2016**, *7*, 519. [[CrossRef](#)]
30. Xue, H.W.; Chen, X.; Mei, Y. Function and regulation of phospholipid signalling in plants. *Biochem. J.* **2009**, *421*, 145–156. [[CrossRef](#)]
31. Divecha, N.; Truong, O.; Hsuan, J.J.; Hinchliffe, K.A.; Irvine, R.F. The cloning and sequence of the C isoform of PtdIns4P 5-kinase. *Biochem. J.* **1995**, *309*, 715–719. [[CrossRef](#)]
32. Mikami, K.; Katagiri, T.; Iuchi, S.; Yamaguchi-Shinozaki, K.; Shinozaki, K. A gene encoding phosphatidylinositol-4-phosphate 5-kinase is induced by water stress and abscisic acid in *Arabidopsis thaliana*. *Plant J. Cell Mol. Biol.* **1998**, *15*, 563–568. [[CrossRef](#)] [[PubMed](#)]
33. Lee, Y.; Kim, Y.W.; Jeon, B.W.; Park, K.Y.; Suh, S.J.; Seo, J.; Kwak, J.M.; Martinoia, E.; Hwang, I.; Lee, Y. Phosphatidylinositol 4,5-bisphosphate is important for stomatal opening. *Plant J. Cell Mol. Biol.* **2007**, *52*, 803–816. [[CrossRef](#)] [[PubMed](#)]
34. Kuroda, R.; Kato, M.; Tsuge, T.; Aoyama, T. *Arabidopsis* phosphatidylinositol 4-phosphate 5-kinase genes PIP5K7, PIP5K8, and PIP5K9 are redundantly involved in root growth adaptation to osmotic stress. *Plant J. Cell Mol. Biol.* **2021**, *106*, 913–927. [[CrossRef](#)] [[PubMed](#)]
35. Liu, X.; Zhai, S.; Zhao, Y.; Sun, B.; Liu, C.; Yang, A.; Zhang, J. Overexpression of the phosphatidylinositol synthase gene (*ZmPIS*) conferring drought stress tolerance by altering membrane lipid composition and increasing ABA synthesis in maize. *Plant Cell Environ.* **2013**, *36*, 1037–1055. [[CrossRef](#)]
36. Lou, Y.; Gou, J.Y.; Xue, H.W. PIP5K9, an *Arabidopsis* phosphatidylinositol monophosphate kinase, interacts with a cytosolic invertase to negatively regulate sugar-mediated root growth. *Plant Cell* **2007**, *19*, 163–181. [[CrossRef](#)]
37. Wen, B.Q.; Xing, M.Q.; Zhang, H.; Dai, C.; Xue, H.W. Rice homeobox transcription factor HOX1a positively regulates gibberellin responses by directly suppressing EL1. *J. Integr. Plant Biol.* **2011**, *53*, 869–878. [[CrossRef](#)]
38. Corratgé-Faillie, C.; Jabnourne, M.; Zimmermann, S.; Véry, A.A.; Fizames, C.; Sentenac, H. Potassium and sodium transport in non-animal cells: The Trk/Ktr/HKT transporter family. *Cell. Mol. Life Sci. CMLS* **2010**, *67*, 2511–2532. [[CrossRef](#)]
39. Han, M.; Wu, W.; Wu, W.H.; Wang, Y. Potassium Transporter KUP7 Is Involved in K(+) Acquisition and Translocation in *Arabidopsis* Root under K(+)-Limited Conditions. *Mol. Plant* **2016**, *9*, 437–446. [[CrossRef](#)]
40. Véry, A.A.; Nieves-Cordones, M.; Daly, M.; Khan, I.; Fizames, C.; Sentenac, H. Molecular biology of K⁺ transport across the plant cell membrane: What do we learn from comparison between plant species? *J. Plant Physiol.* **2014**, *171*, 748–769. [[CrossRef](#)]
41. Li, W.; Xu, G.; Alli, A.; Yu, L. Plant HAK/KUP/KT K transporters: Function and regulation. *Semin. Cell Dev. Biol.* **2018**, *74*, 133–141. [[CrossRef](#)]
42. Qin, Y.J.; Wu, W.H.; Wang, Y. *ZmHAK5* and *ZmHAK1* function in K uptake and distribution in maize under low K conditions. *J. Integr. Plant Biol.* **2019**, *61*, 691–705. [[CrossRef](#)] [[PubMed](#)]
43. Conesa, A.; Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Talon, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [[CrossRef](#)] [[PubMed](#)]
44. Ye, J. WEGO: A web tool for plotting GO annotations. *Nucleic Acids Res.* **2006**, *34*, W293–W297. [[CrossRef](#)] [[PubMed](#)]
45. Li, J.; Harata-Lee, Y.; Denton, M.D.; Feng, Q.; Rathjen, J.R.; Qu, Z.; Adelson, D.L. Long read reference genome-free reconstruction of a full-length transcriptome from *Astragalus* membranaceus reveals transcript variants involved in bioactive compound biosynthesis. *Cell Discov.* **2017**, *3*, 17031. [[CrossRef](#)] [[PubMed](#)]
46. Alamancos, G.P.; Pagès, A.; Trincado, J.L.; Bellora, N.; Eyra, E. Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA* **2015**, *21*, 1521–1531. [[CrossRef](#)] [[PubMed](#)]
47. Sun, L.; Luo, H.; Bu, D.; Zhao, G.; Yu, K.; Zhang, C.; Liu, Y.; Chen, R.; Zhao, Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* **2013**, *41*, e166. [[CrossRef](#)]
48. Kong, L.; Yong, Z.; Ye, Z.Q.; Liu, X.Q.; Ge, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* **2007**, *35*, W345–W349. [[CrossRef](#)]
49. Finn, R.D.; Penelope, C.; Eberhardt, R.Y.; Eddy, S.R.; Jaina, M.; Mitchell, A.L.; Potter, S.C.; Marco, P.; Matloob, Q.; Amaia, S.V. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [[CrossRef](#)]
50. Li, A.; Zhang, J.; Zhou, Z. PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinform.* **2014**, *15*, 311. [[CrossRef](#)]
51. Dewey, C.N.; Li, B. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323.
52. Kim, G.B.; Gao, Y.; Palsson, B.O.; Lee, S.Y. DeepTFactor: A deep learning-based tool for the prediction of transcription factors. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2021171118. [[CrossRef](#)] [[PubMed](#)]
53. Yi, Z.; Chen, J.; Sun, H.; Rosli, H.G.; Pombo, M.A.; Zhang, P.; Banf, M. iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* **2016**, *9*, 1667–1670.
54. Eddy, S.R.; Pearson, W.R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **2011**, *7*, e1002195. [[CrossRef](#)] [[PubMed](#)]

55. Finn, R.D.; Alex, B.; Jody, C.; Penelope, C.; Eberhardt, R.Y.; Eddy, S.R.; Andreas, H.; Kirstie, H.; Liisa, H.; Jaina, M. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)]
56. Duvaud, S.; Gabella, C.; Lisacek, F.; Stockinger, H.; Durinx, C. ExPASy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res.* **2021**, *49*, W216–W227. [[CrossRef](#)]
57. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
58. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Higgins, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
59. Sudhir, K.; Glen, S.; Koichiro, T. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874.
60. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
61. Liu, Q.; Qi, X.; Yan, H.; Huang, L.; Nie, G.; Zhang, X. Reference Gene Selection for Quantitative Real-Time Reverse-Transcriptase PCR in Annual Ryegrass (*Lolium multiflorum*) Subjected to Various Abiotic Stresses. *Molecules* **2018**, *23*, 172. [[CrossRef](#)]