International Journal of
Molecular Sciences

MDPI

Article

A Comprehensive Identification and Expression Analysis of VQ Motif-Containing Proteins in Sugarcane (*Saccharum spontaneum* L.) under Phytohormone Treatment and Cold Stress

Ying Liu [1,2], Xiaolan Liu [1,2], Dandan Yang [1], Ze Yin [1], Yaolan Jiang [1], Hui Ling [3], Ning Huang [3], Dawei Zhang [1,2], Jinfeng Wu [1,2], Lili Liu [1,2], Liping Xu [4], Mingli Yan [1,5], Youxiong Que [4,*] and Dinggang Zhou [1,2,4,*]

1 Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization, School of Life Science, Hunan University of Science and Technology, Xiangtan 411201, China; ly1071457534@163.com (Y.L.); liuxiaolan1018@126.com (X.L.); ydd2595836963@163.com (D.Y.); yz18390246279@163.com (Z.Y.); jyl0480_17@163.com (Y.J.); zhangdawei.hnust@foxmail.com (D.Z.); wujinfeng.hnust@foxmail.com (J.W.); lilyliu@hnust.edu.cn (L.L.); ymljack@126.com (M.Y.)
2 Key Laboratory of Integrated Management of the Pests and Diseases on Horticultural Crops in Hunan Province/Key Laboratory of Ecological Remediation and Safe Utilization of Heavy Metal-Polluted Soils, College of Hunan Province, Xiangtan 411201, China
3 College of Agriculture, Yulin Normal University, Yulin 537000, China; linghuich@ylu.edu.cn (H.L.); hningch@ylu.edu.cn (N.H.)
4 Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture/National Engineering Research Center for Sugarcane, Ministry of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China; xlpmail@126.com
5 Crop Research Institute, Hunan Academy of Agricultural Sciences, Changsha 410000, China
* Correspondence: queyouxiong@fafu.edu.cn (Y.Q.); dgzhoucn@hnust.edu.cn (D.Z.); Tel./Fax: +86-731-5829 (ext. 1416) (D.Z.)

Abstract: The VQ motif-containing proteins play a vital role in various processes such as growth, resistance to biotic and abiotic stresses and development. However, there is currently no report on the *VQ* genes in sugarcane (*Saccharum* spp.). Herein, 78 *VQ* genes in *Saccharum spontaneum* were identified and classified into nine subgroups (I-IX) by comparative genomic analyses. Each subgroup had a similar structural and conservative motif. These *VQ* genes expanded mainly through whole-genome segmental duplication. The cis-regulatory elements (CREs) of the *VQ* genes were widely involved in stress responses, phytohormone responses and physiological regulation. The RNA-seq data showed that *SsVQ* gene expression patterns in 10 different samples, including different developmental stages, revealed distinct temporal and spatial patterns. A total of 23 *SsVQ* genes were expressed in all tissues, whereas 13 *SsVQ* genes were not expressed in any tissues. Sequence Read Archive (SRA) data showed that the majority of *SsVQs* responded to cold and drought stress. In addition, quantitative real-time PCR analysis showed that the *SsVQs* were variously expressed under salicylic acid (SA), jasmonic acid (JA), abscisic acid (ABA) and cold treatment. This study conducted a full-scale analysis of the *VQ* gene family in sugarcane, which could be beneficial for the functional characterization of sugarcane *VQ* genes and provide candidate genes for molecular resistance breeding in cultivated sugarcane in the future.

Keywords: *Saccharum spontaneum*; *VQ* gene; genome-wide analysis; Transcriptome; phytohormone; cold stress

check for updates

Citation: Liu, Y.; Liu, X.; Yang, D.; Yin, Z.; Jiang, Y.; Ling, H.; Huang, N.; Zhang, D.; Wu, J.; Liu, L.; et al. A Comprehensive Identification and Expression Analysis of VQ Motif-Containing Proteins in Sugarcane (*Saccharum spontaneum* L.) under Phytohormone Treatment and Cold Stress. *Int. J. Mol. Sci.* **2022**, *23*, 6334. https://doi.org/10.3390/ijms23116334

Academic Editor: Marcello Iriti

Received: 10 May 2022
Accepted: 3 June 2022
Published: 6 June 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

Sugarcane (*Saccharum* spp. hybrid) is the most important crop for sugar-producing and raw material for fuel alcohol around the world [1,2]. It contributes to over 80% of the sugar and approximately 40% of the ethanol global production [1,3]. The most modern sugarcane

varieties have been derived from hybrids between *S. spontaneum* and *S. officinarum* [4]. *S. officinarum* (called "noble cane") has higher sugar content, while *S. spontaneum* contributes with disease resistance genes and ratooning ability [5]. Various stresses, including pathogen infection, drought, pest attack, cold and salinity stresses, severely and consistently reduce sugar production [6–9]. Basnayake et al. reported that drought had reduced sugarcane production by up to 60% [10]. Mosaic, one of the main viral sugarcane diseases, has resulted in stunted plants, reduced photosynthesis, destruction of chlorophyll and inhibition of plant growth and which can ultimately reduce sugarcane yield by 10–50%, or even 60–80% [11,12]. Approximately 1300 insect pests have been accounted for attacking sugarcane around the world, and stem borer is the most harmful pest in most countries, resulting in yield losses of nearly 25–30% [1,13,14]. Like all plants, sugarcane has evolved fancy mechanisms to respond to external stimuli [15]. These responses formed sophisticated signaling pathways, involving many endogenous and exogenous signal molecules, such as phytohormones and other regulators, which systematically modulate transcriptional and post-transcriptional processes [15,16]. To understand the sugarcane growth and development and response to environmental stressors, it is significant to investigate the complexity of transcription factors (TFs) and transcription-associated proteins (TAPs) present in the genome and the functions [4,17].

One of the largest transcriptional regulator families in plants is the VQ proteins that work solely or in combination with other TFs to regulate the various life processes, including plant growth and development and responses to different stress [18–20]. The VQ proteins, named by the highly conserved amino acid motif (FxxVQxhTG), have been widely found in monocotyledons and dicotyledons [19,21–27]. In addition, the types of VQ proteins are varied among plants. For example, there are four types in rice (LTG, FTG, VTG and ITG) and six types in tobacco (LTG, FTG, YTG, VTG, LTA and LTV) [22,28]. Though previously thought to be plant-specific TFs, VQ proteins have also been found in bacteria, animals and fungi, implying that the *VQ* gene family has a common ancient ancestor [29].

VQ proteins play a vital regulatory role in seed development, as well as vegetative plant growth, biotic and abiotic stress responses. *IKU1* (*AtVQ14*/AT2G35230) is involved in regulating endosperm development and seed growth [30]. *AtVQ22* could alleviate JA stress through interaction with *AtWRKY28* and *AtWRKY51* [31]. Dong et al. reported that overexpression of apple *VQ37* could affect vegetative growth and reproductive growth in *Arabidopsis* and tobacco plants [27]. The *AtVQ15* plays a negative role in response to osmotic stress [32]. Meanwhile, increasingly more studies have shown that many VQ proteins interact with WRKY TFs in response to abiotic and biotic stress. *MaWRKY26* can physically interact with *MaVQ5* to control the JA biosynthetic [20]. *OsVQ7* could improve the tolerance of various stresses and development through interaction with *OsWRKY24* [33].

At present, *VQ* genes have been identified and analyzed in many plants, including both monocotyledons and dicotyledons. *Arabidopsis, Oryza sativa* and *Zea mays* were found to have 34, 39 and 61 *VQ* genes, respectively [19,25,28]. However, there is little available information about the VQ protein gene family in sugarcane. In this study, we carried out a comprehensive analysis of the *VQ* gene family based on *S. spontaneum* AP85–441 which has been sequenced and assembled. In the present study, 78 *SsVQs* in *S. spontaneum* were identified. Their classification, gene structure, conserved domains, motif composition, chromosomal location distribution, evolution and CREs were analyzed. In addition, the expression patterns for *VQ* genes in 10 different sugarcane tissues, cold and drought stress, and under hormone (ABA, SA and JA) treatment and cold stress were investigated using RNA-seq data, SRA data and real-time quantitative polymerase chain reaction (RT-qPCR), respectively. The present study aimed to promote the functional study of *VQ* genes in sugarcane, especially for the understanding of the mechanism of *VQ* genes under phytohormone and cold stress. Moreover, our results could contribute to providing key candidate genes for molecular stress resistance breeding in sugarcane.

2. Results

2.1. Identification and Sequence Analysis of SsVQ Genes

In total, we identified 78 VQ motif-containing proteins using BlastP and HMMER software and named them from *SsVQ1* to *SsVQ78* based on their physical locations on the chromosomes. The total number of *VQ* genes in sugarcane was larger than that in *Arabidopsis* (34), rice (40) and maize (61). ExPasy prediction revealed that these 78 VQ proteins have different physical and chemical properties. The amino acid lengths ranged from 77 aa (SsVQ15) to 812 aa (SsVQ30), with an average of 258 aa and most of them were less than 300 aa (Table S1). The molecular weights ranged from 8063.42 Da (SsVQ15) to 85,904.61 Da (SsVQ30) and their isoelectric points ranged from 4.73 (SsVQ74, SsVQ75 and SsVQ76) to 11.59 (SsVQ11). To investigate the protein hydrophobicity, the GRAVY score was conducted, and the results showed that the GRAVY scores all were negative except for the SsVQ29, indicating that most SsVQ proteins are hydrophilic. Moreover, the subcellular localization prediction showed that 44 *SsVQ* genes were located in the nucleus, 15 *SsVQ* genes were located in the chloroplasts, 13 *SsVQ* genes were located in the mitochondrion, four *SsVQ* genes (*SsVQ64*, *SsVQ74*, *SsVQ75* and *SsVQ76*) were located in the cytoplasm and two *SsVQ* genes (*SsVQ37* and *SsVQ38*) were located in the endoplasmic reticulum (Table S1).

2.2. Phylogenetic Analysis and Multiple Sequence Alignment

To investigate the relationships of *SsVQ* genes among sugarcane, maize, rice and *Arabidopsis*, we conducted a phylogenetic tree based on their protein sequences (Figure 1). We found that maize has a closer relationship with sugarcane than rice and *Arabidopsis*. Based on the relationship with *AtVQs* and *OsVQs* and the domains of SsVQ protein, they were divided into nine groups, named Group I–IX. For the 78 SsVQ proteins, GroupVI contains three VQ proteins; Group VII has the largest amount, with 19 VQ proteins. Group I, II, III, IV, V, VIII and IX contain 15, 8, 6, 9, 4, 6 and 8 members, respectively.

At the same time, we performed the multiple sequence alignment and found five types of VQ specificity domain (Figure 2). The results showed that all identified SsVQ proteins contained the motif FxxhVQxhTG, while the x represents any amino acid and h represents a hydrophobic amino acid. 59/78 FxxxVQxLTG, 12/78 FxxxVQxFTG, 2/78 FxxxVQxVTG (*SsVQ6* and *SsVQ18*), 2/78 FxxxVQxITG (*SsVQ21* and *SsVQ26*), 3/78 FxxxVHQxVTG (*SsVQ66*, *SsVQ67* and *SsVQ68*). Different types of VQ domains indicated that they might have different biological functions.

2.3. Conserved Motifs and Gene Structures of the VQ Gene Family

To predict the function of *VQ* genes, we detected the conserved motifs. The results indicated that these 78 *SsVQs* contained 10 conserved motifs and the motif length ranged from 11 aa to 50 aa (Figure 3; Table S2). All VQ proteins contain motif 1, which was VQ domain and each *SsVQ* member contains 1–6 conserved motifs. Moreover, an unrooted phylogenetic tree was constructed based on VQ protein sequences, suggesting that the motif classification of *VQ* genes was consistent with the phylogenetic tree. We found that most groups possess three motifs, which suggested that every group might have special functions with a highly conserved amino acid residue, while seven VQ proteins only have one VQ motif (VQ1, VQ15, VQ27, VQ34, VQ45, VQ47 and VQ64). Notably, some motifs appeared in one sub-branch of *SsVQ*. For example, motif 2 only exists in Group III. Motif 8 only exists in Group V. Through the *VQ* gene structures analysis, only three of the *SsVQs* have UTR. Interestingly, 73.08% (57/78) of *SsVQ* genes are intronless genes, which indicated that many introns might be lost during the stage of *VQ* gene evolution.

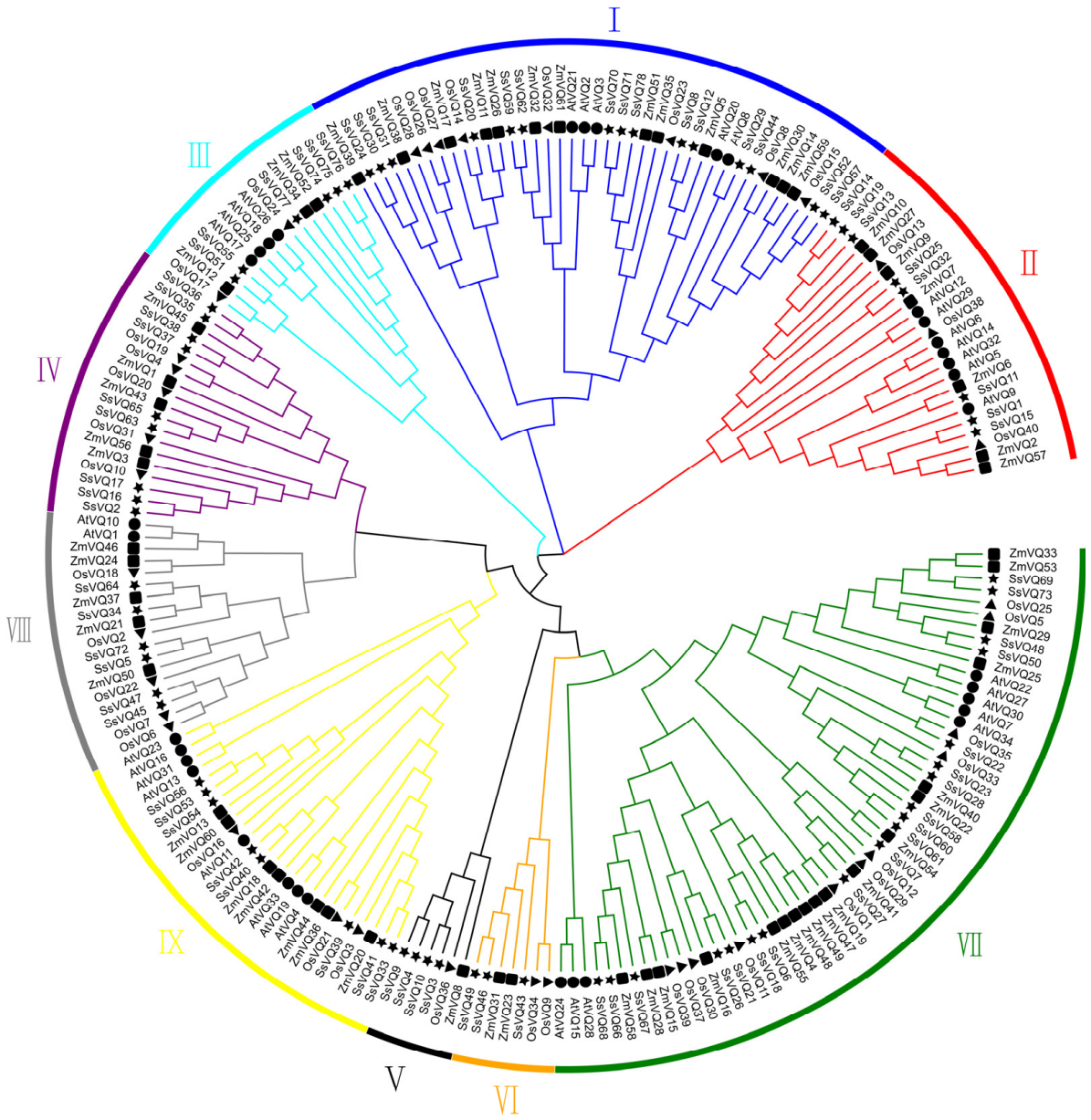

Figure 1. Phylogenetic analysis and family classification of the VQ domains. The NJ tree was constructed using MEGA X with 1000 boosted replicates. The different colored arcs indicate different groups of the VQ domain. Proteins from AP85-441 (*S. spontaneum*), maize (*Zea mays*), rice (*O. sativa*) and *Arabidopsis* are represented by stars, squares, triangles and circles, respectively.

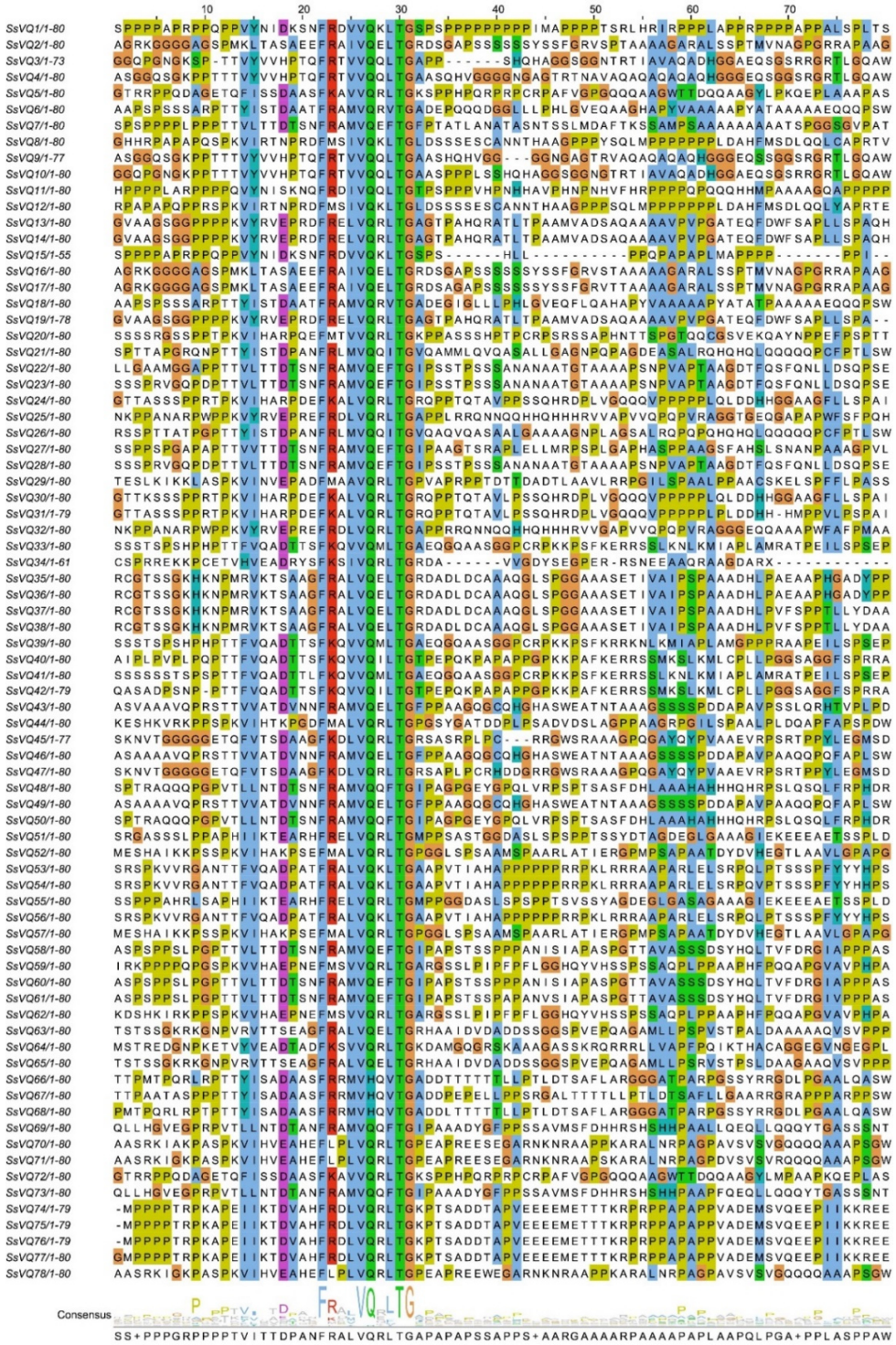

Figure 2. Multiple sequence alignment of the VQ proteins in *S. spontaneum*. The sequences were aligned using the Jalview software. The highly conserved motif is FxxxVQxLTG.

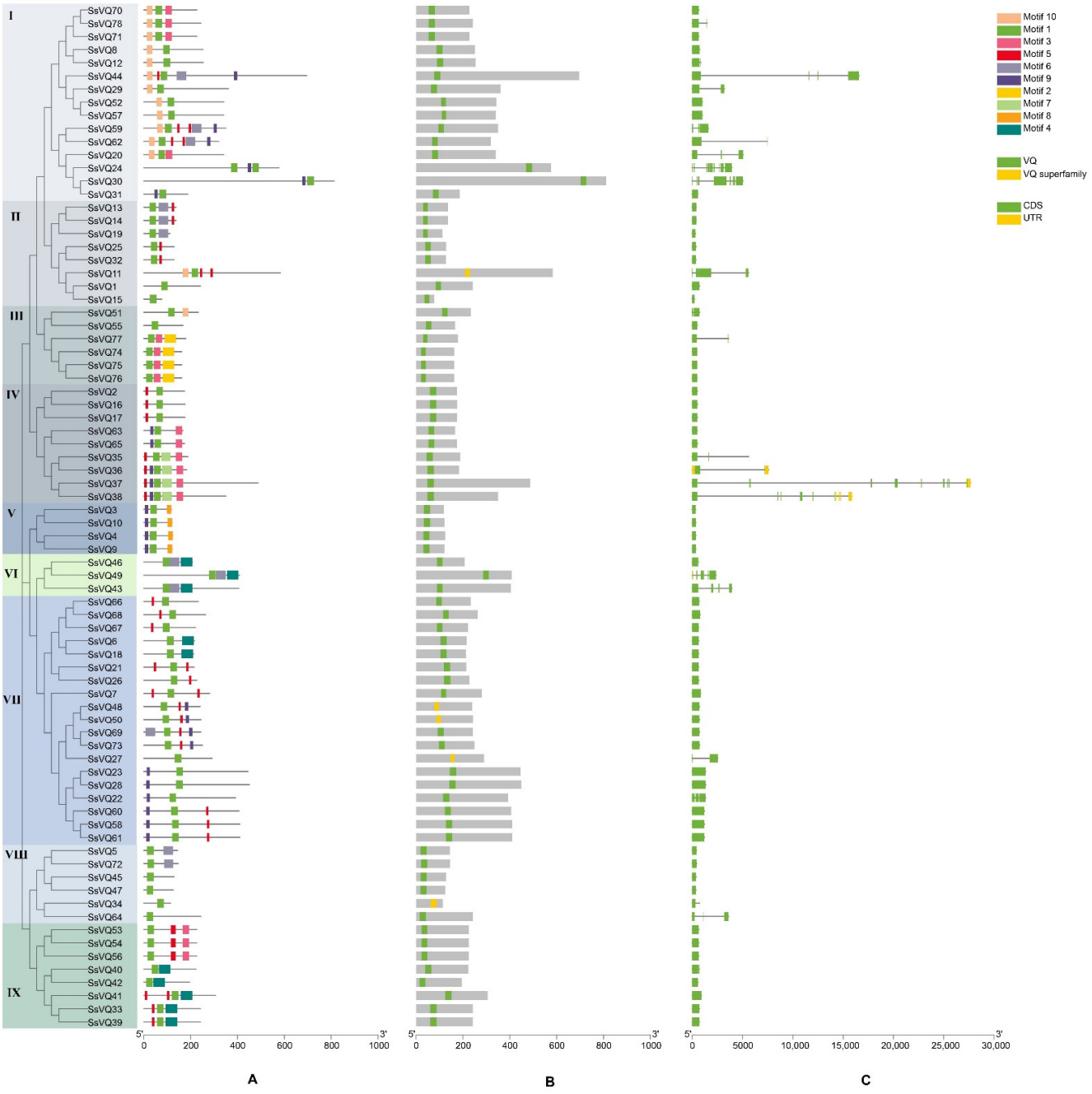

Figure 3. Phylogenetic tree, conserved motif, conserved domain, and the gene structure of the *VQ* genes. The phylogenetic tree was constructed using MEGA X software based on VQ protein sequences. (**A**) Conserved motifs of the VQ proteins. Each motif is represented with a specific color. (**B**) The conserved domain of VQ protein. The green and yellow boxes represent the VQ domain. (**C**) Gene structure of *VQ* gene. The untranslated 5′- and 3′-regions, introns and exons are represented with yellow box, green box and black lines, respectively.

2.4. Chromosome Distribution and Gene Duplication Analysis

To understand the position of *SsVQ* genes on the chromosome, we conducted a chromosomal location map of these *SsVQ* genes (Figure 4). *SsVQs* are located on all sugarcane chromosomes, except chromosome 7C. A large number of *SsVQs* are located on the two ends of chromosomes. Specifically, chromosome 1B and chromosome 1D all contained six genes. Segmental or tandem duplicates in many gene families are the main expanding pattern in plants [34]. To better study the evolution of *SsVQ* genes, we

further predicted the gene duplication events by the MCScanX software (Figure 5). We found that 24 pairs of *VQ* genes originated from segmental duplication, and three pairs of genes (*SsVQ3*/*SsVQ4*, *SsVQ9*/*SsVQ10* and *SsVQ67*/*SsVQ68*) were involved in tandem duplication events. The Ka and Ks rates and the Ka/Ks of these *VQ* gene pairs were calculated. Results showed that 18 pairs were <1.0, two pairs were >1.0, and the other four did not have values (Table S3). Therefore, purifying selection might be the primary pressure during the evolutionary period of *SsVQs*. We further detected the collinear relationship among *Arabidopsis*, rice, *Sorghum bicolor* and sugarcane. The results showed that sugarcane and sorghum have a closer relationship (Figure S1).

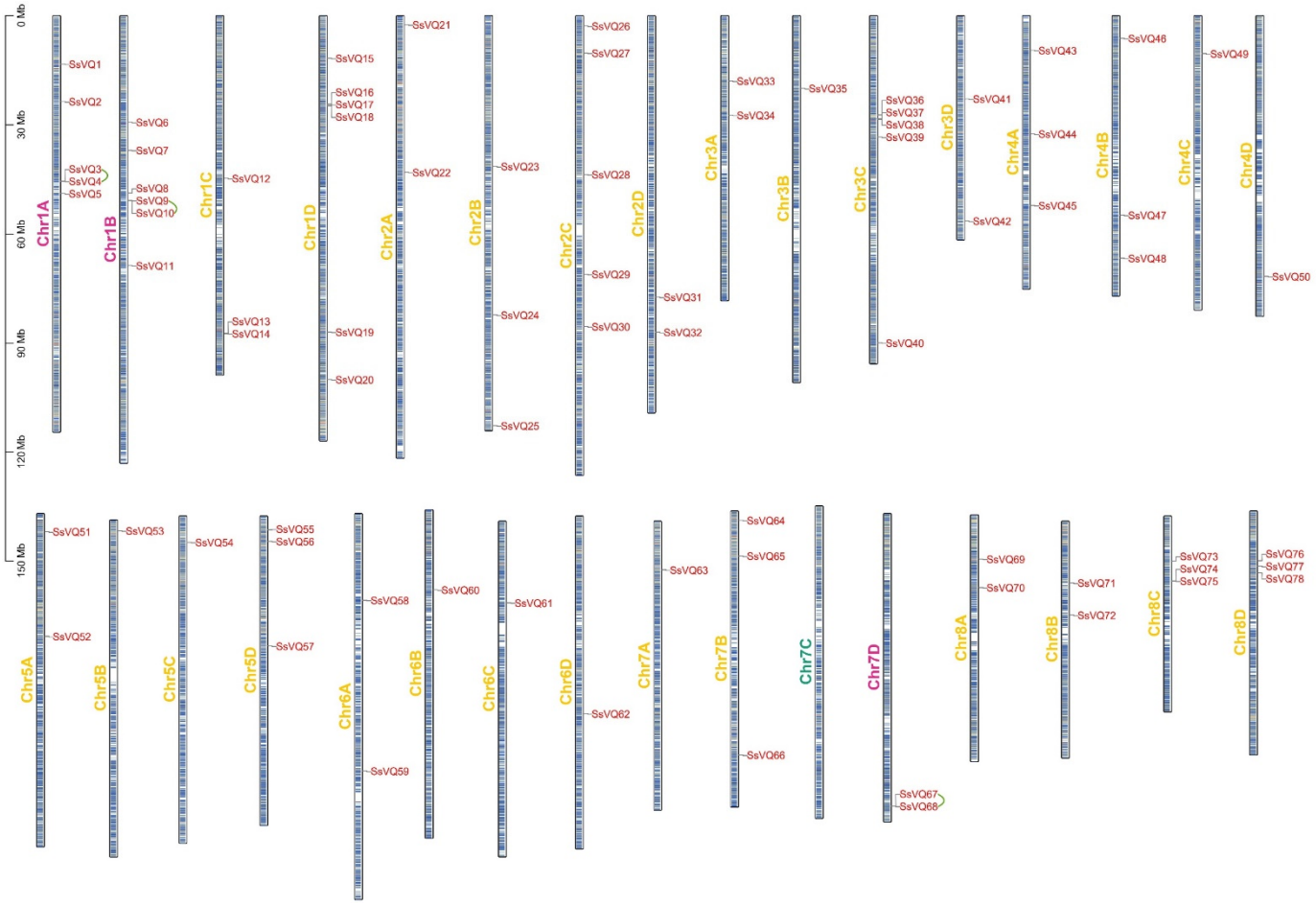

Figure 4. Schematic representations for the chromosomal distribution of *S. spontaneum VQ* genes. A green line between the two gene names indicated that they were tandem repeat gene pairs. The chromosome number was indicated to the left of each chromosome. *SsVQ* gene numbers are shown on the right of each chromosome. Scale bar on the left indicates the chromosome lengths (Mb).

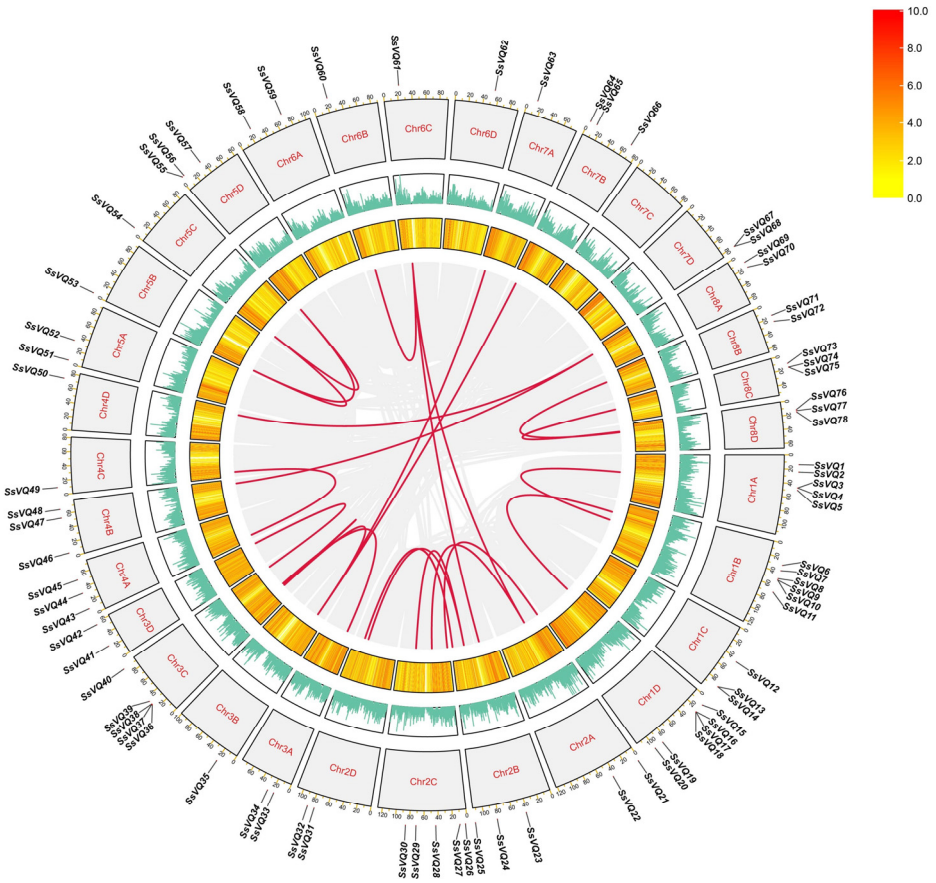

Figure 5. Schematic representations highlighting the interchromosomal relationships of the *SsVQ* genes. Gray lines indicated all syntenic blocks in the sugarcane genome, and the red lines indicated duplicated *VQ* gene pairs.

2.5. Identification of CREs in the Promoter Regions of SsVQ Genes

In this study, we identified putative cis-elements of 3000 bp located on *SsVQ* genes to predict the possible gene function by using PlantCARE (Table S4). The results showed that the numerous CREs widely distributed in the promoter region of *SsVQ* genes including stress responsiveness, phytohormone responsiveness, light responsiveness and growth and development. The light, ABA and MeJA responsive elements were the most numerous in the promoter region of *SsVQs*. Among them, light responsive elements were found in all members. In total, 77/78 *SsVQs* contained ABA response elements except the *SsVQ2*, and 76/78 *SsVQs* possessed MeJA response elements except the *SsVQ2* and *SsVQ15*, while 73/78 *SsVQs* contained anaerobic induction elements (ARE). Moreover, 53 (67.95%) *SsVQs* contained drought responsive elements (MBS), 52 (66.67%) *SsVQs* contained gibberellin-responsive elements (TATC-box, P-box and GARE-motif), 43 (55.13%) *SsVQs* contained SA responsive elements, 42 (53.85%) *SsVQs* contained auxin-responsive elements (TGA-element and AuxRR-core) and 38 (48.72%) *SsVQs* contained low-temperature response elements (LTR), while 34 (43.59%) *SsVQs* contained defense and stress responsiveness (TC-rich repeats). In addition, the circadian control element, GCN4_motif (involved in endosperm expression), circadian control element and GC-motif (involved in anoxic specific inducibility) were also identified. These results indicated that most of the *SsVQ* genes might respond to various plant biotic and abiotic stresses and play a vital role in sugarcane growth and development.

2.6. Protein-Protein Network of the VQ Proteins

Based on the STRING database with maize as a reference, a PPI network was constructed to explore the physical and functional association (Figure 6). Some of these proteins

(GRMZM2G174650_P01-GRMZM2G023921_P01, GRMZM2G174650_P01-GRMZM2G059064_P01, GRMZM2G174650_P01-GRMZM2G118172_P01) have strong interaction. These results could provide a piece of useful information for subsequent studies on the regulatory network of VQs.

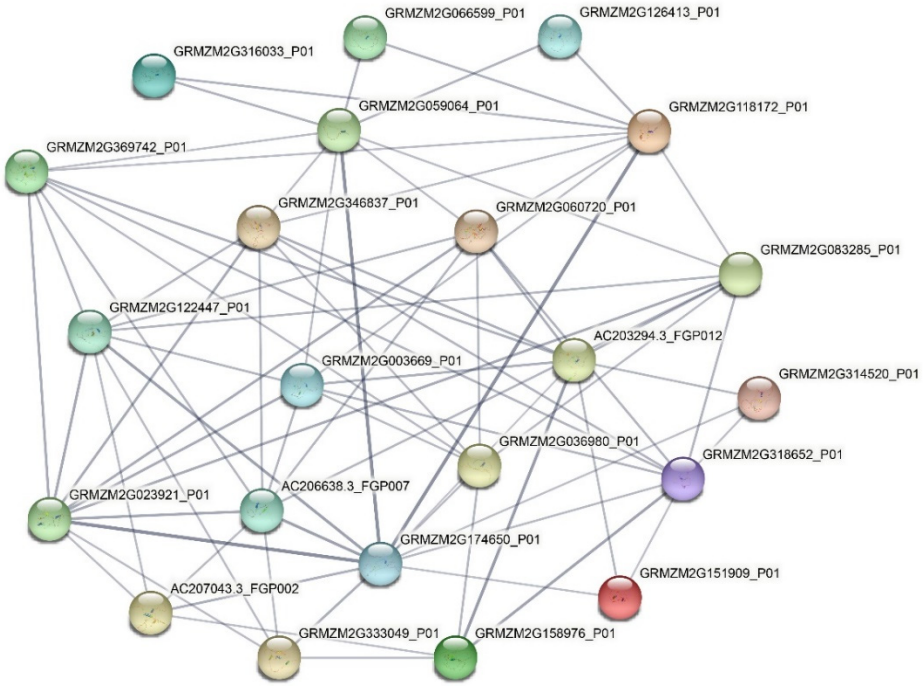

Figure 6. Schematic representation of protein–protein interaction (PPI) networks of the SsVQ proteins based on their orthologs in maize. The detailed information of the network is shown in Table S5. Gray lines connect proteins within the PPI networks with darker colors and thicker lines indicating higher core PPI values.

2.7. Expression Pattern for SsVQ Genes during Sugarcane Development Based on RNA-Seq Data

Using the public RNA-seq data, we analyzed the *SsVQ* gene expression profiles in 10 different tissues to understand spatiotemporal expression patterns (Figure 7). In general, most *SsVQs* are expressed during the stage of sugarcane growth and development but with different expression. Among the *SsVQ* genes, 13 were not expressed in all tissues, while 23 were expressed in all 10 tissues (FPKM > 0). Most of these *SsVQ* genes were found expressed in more than one detected organ. *SsVQ73* and *SsVQ69* had similar expression patterns and were highly expressed at the early and mature stages of leaf and steam, while *SsVQ37* and *SsVQ38* were highly expressed at the stage of seedling–steam and seedling–leaf. In addition, *SsVQ2*, *SsVQ16*, *SsVQ33*, *SsVQ257* and *SsVQ66* were higher expressed in the stage of leaf development than the other nine stages, indicating that they are related to the growth and development of corresponding tissues. These results implied that *SsVQs* may play an important role in the growth and development of sugarcane, but with individual functional modes.

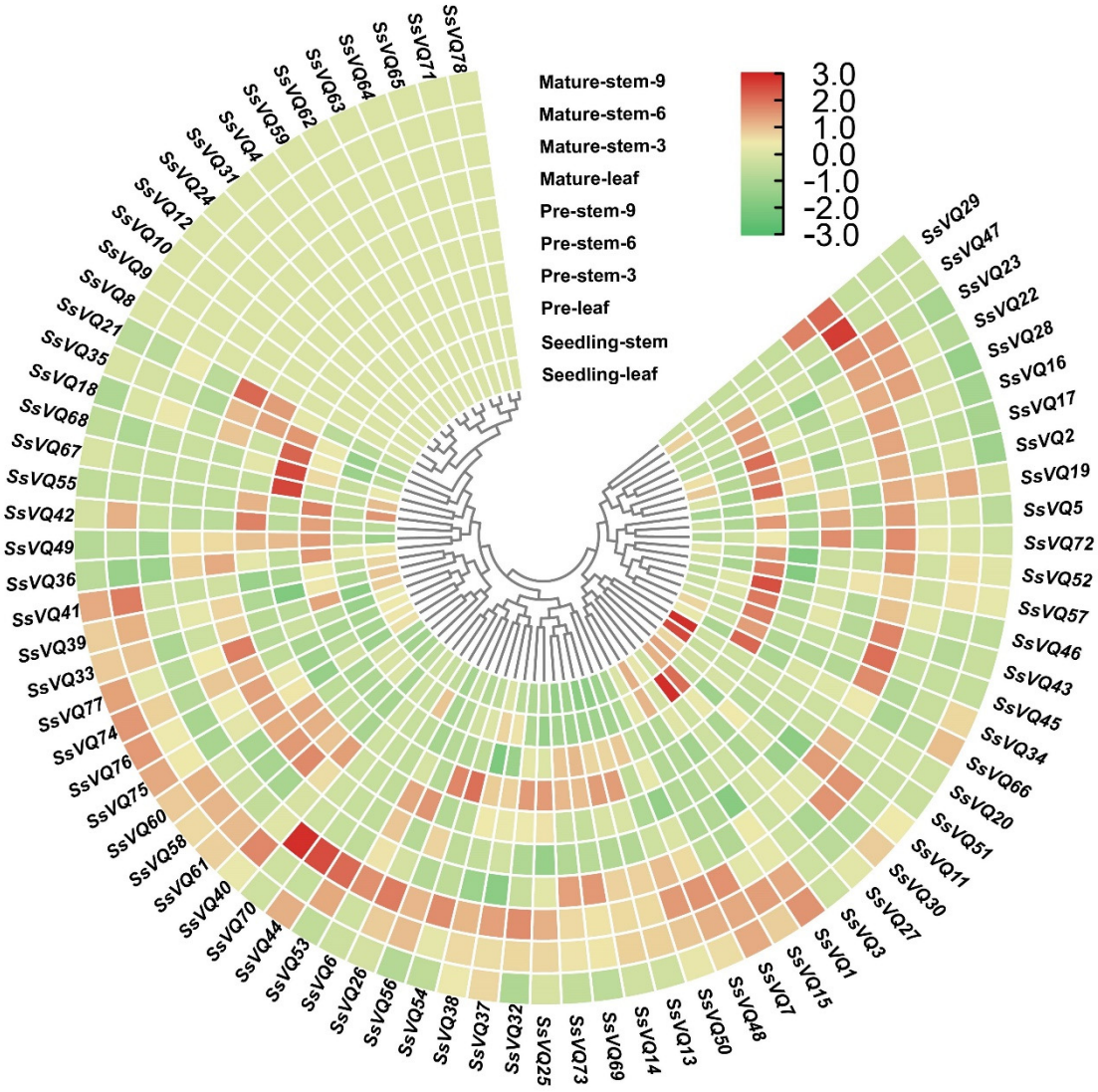

Figure 7. Expression profiles of the *VQ* genes from *Saccharum spontaneum* at different stages. The color bar represents the normalized values ($\log_2$ FPKM). The original normalized values are shown in Table S6.

2.8. Expression Pattern for SsVQ Genes under Cold and Drought Treatments Based on SRA Data

Since cis-elements responding to drought and cold stress widely existed in the promoters of the *SsVQ* genes. We analyzed the expression patterns for *SsVQs* under drought and cold stress to further investigate the potential functions of *SsVQ* genes (Figure 8). Under drought stress, the *SsVQs* had different expression profiles. For example, six *SsVQs* (*SsVQ20*, *SsVQ24*, *SsVQ31*, *SsVQ59*, *SsVQ62* and *SsVQ65*) were not expressed. Most *SsVQ* genes had the highest expression levels after 10 days of recovery, especially *SsVQ4*, *SsVQ9*, *SsVQ33*, *SsVQ35*, *SsVQ39*, *SsVQ69* and *SsVQ73*. While *SsVQ40*, *SsVQ58*, *SsVQ60* and *SsVQ61* had the lowest expression levels after 10 days recovery. The expression level of the *SsVQ38* gene showed a continuous downward trend after drought treatment (2, 6 and 10 days drought) and after 10 days of recovery treatment. As for cold treatment, 15 *SsVQs* were not expressed. Different from the drought treatment, the expression levels of most *SsVQs* did not change significantly but remained relatively stable. Three of the *SsVQs* (*SsVQ38*, *SsVQ69* and *SsVQ73*) consistently had the highest expression level during the treatment. These results showed that *SsVQs* may play different roles in response to drought and cold stress.

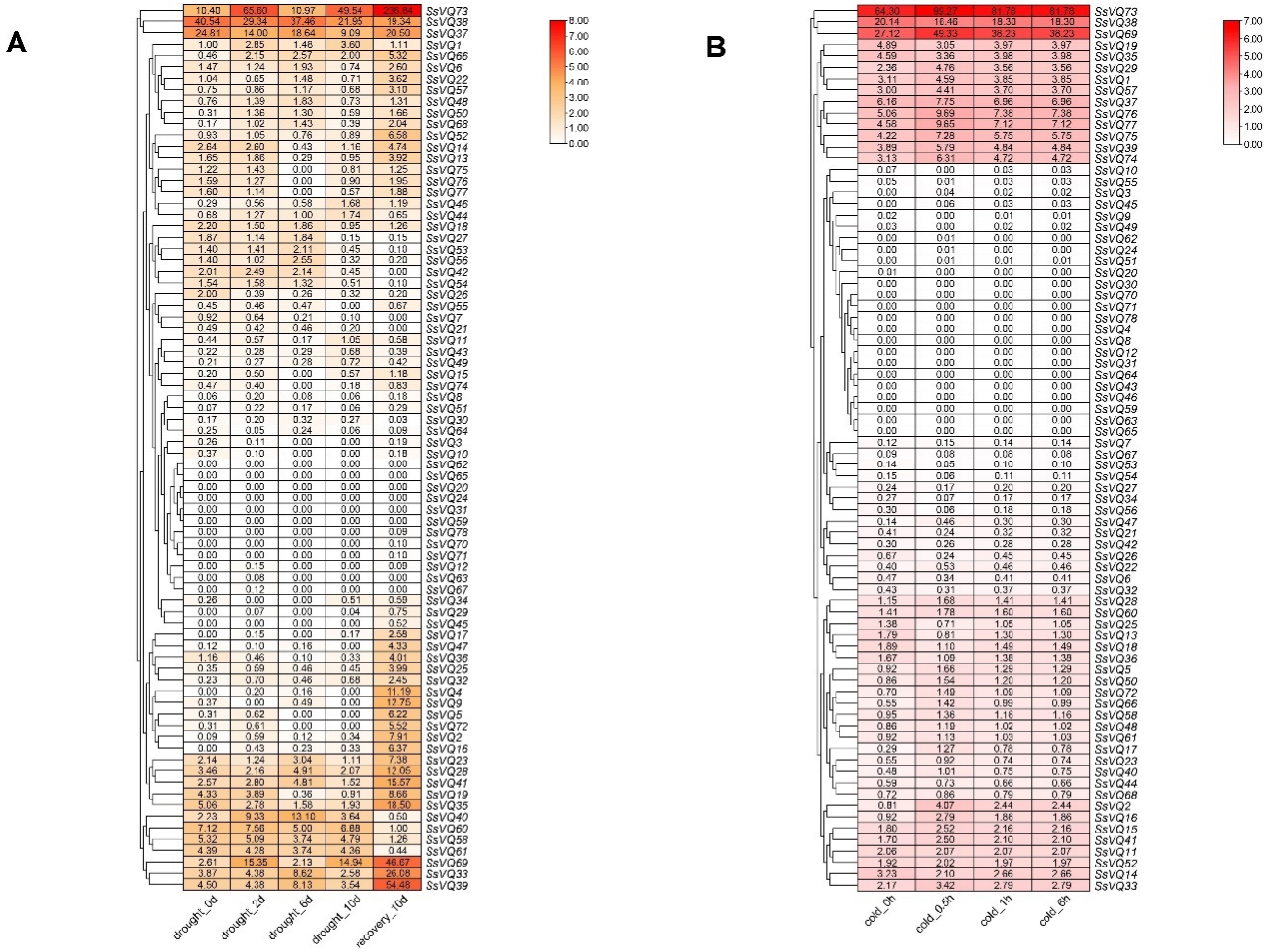

Figure 8. Expression profiles of *SsVQs* in response to drought and cold stresses. The number shown in the box is the original trimmed mean of M-values (TMM), which represent the expression levels of *SsVQs*. The color bar represented the normalized values (log $_2$ TMM). (**A**) The heatmap of *SsVQ* genes under drought stress. (**B**) The heatmap of *SsVQ* genes under cold stress.

2.9. Expression Analyses of VQ genes in Response to Hormone Treatments Using qRT-PCR

To further determine whether the *VQ* gene expression patterns were affected by hormone treatments, we examined seven *VQ* genes and performed a qRT-PCR to analyze the expression patterns under different treatments (Figure 9). SA, JA and ABA treatment could affect the expression of all seven genes. After 12 h of treatment, the up-regulation multiple for most *VQ* genes was the highest among all treatment time points compared with the control group. Different treatments had different effects on gene expression. Upon ABA treatment, four genes (*VQ8*, *VQ30*, *VQ73* and *VQ76*) were up-regulated among all time points. Two genes (*VQ27* and *VQ44*) showed a tendency towards down-regulation in all time points. Under SA stress, *VQ27* and *VQ57* were down-regulated after 1 h of treatment and had few differences after 6 h and 12 h of treatment. In the case of JA treatment, *VQ73* was rapidly up-regulated after 1 h of treatment. Interestingly, *VQ8* was dramatically up-regulated among all hormone treatments. The expression pattern for *VQ76* was the same under the three hormones treatment, which increased first and then decreased, and the expression level was the highest at 6 h. The results indicated that *SsVQs* may play an important role in response hormones.

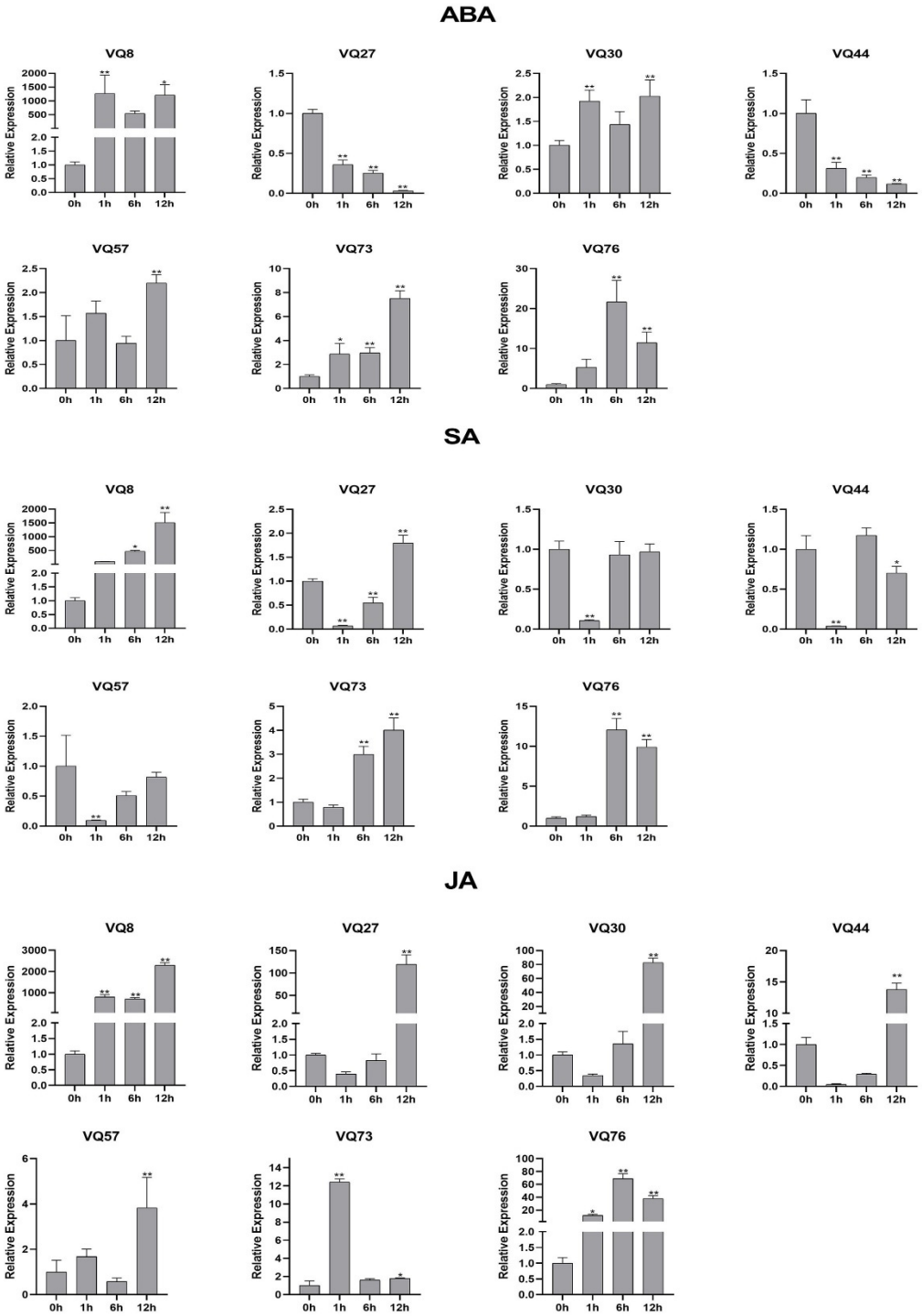

Figure 9. The expression profiles of seven selected *VQ* genes in response to different phytohormones. Plants were treated with ABA (100 μmol/L), SA (5 mmol/L) and MeJA (100 μmol/L) for 0, 1, 6 and 12 h. The 25s-RNA gene was used as the internal control and to normalize expression data. Relative transcript abundance was normalized relative to the expression of CK at 0 h. The $2^{-\Delta\Delta Ct}$ method was used to calculate the expression of target genes at different times and treatments. Error bars represented the standard deviation of the mean. * means $p < 0.05$, ** means $p < 0.01$.

2.10. SsVQ Gene Expression following Abiotic Treatments Using qRT-PCR

The above data suggested that *SsVQ* gene expression was affected by hormone treatment; we also investigated their expression levels under the cold environment (Figure 10). The expression trend for the three genes (*VQ44*, *VQ73* and *VQ76*) was the same, which

increased first after 1 h of treatment and then decreased continuously after 6 h and 12 h of treatment. However, the expression of *VQ8* decreased under cold stress, which was different from that under hormone treatment. The expression pattern for *VQ30* and *VQ57* was continuously up-regulated among all treatment time points.

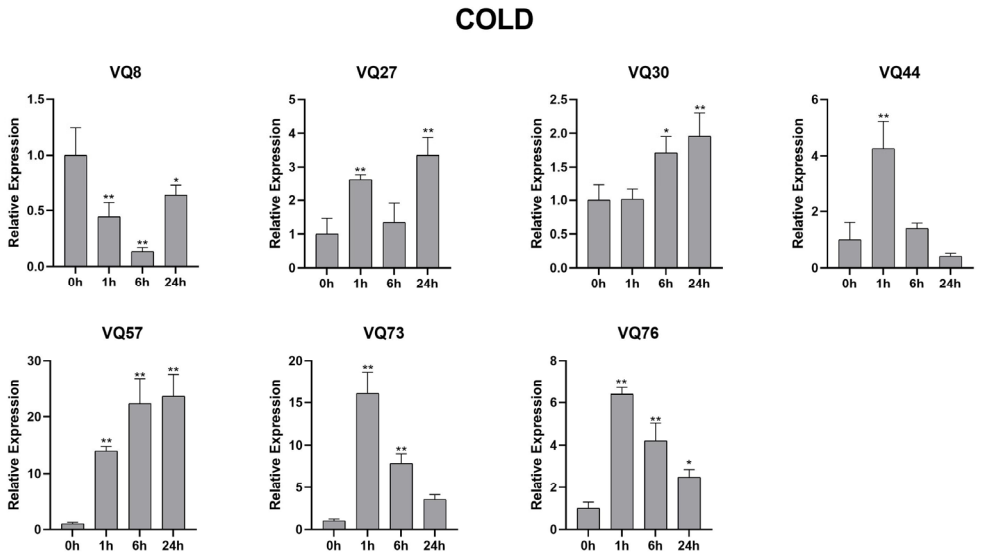

Figure 10. The expression profiles of seven selected *VQ* genes in response to cold stress. Plants were treated with cold (4 °C) stress for 0, 1, 6 and 24 h. The 25s-RNA gene was used as the internal control and to normalize expression data. Relative transcript abundance was normalized relative to the expression of CK at 0 h. The $2^{-\Delta\Delta Ct}$ method was used to calculate the expression of target genes at different times and treatments. Error bars represented the standard deviation of the mean. * means $p < 0.05$, ** means $p < 0.01$.

3. Discussion

VQ protein, a type of plant specific protein, is involved in plant development and can respond to different stresses [25,27,35]. Hence, we conducted a completed genome-wide analysis of sugarcane VQ proteins using bioinformatics analysis and qRT-PCR to understand their regulation with the environment changes.

In this study, 78 *SsVQ* genes were identified, which was higher than that in rice (39), maize (61) and *Arabidopsis* (34). The genome size varies in different plants, rice (466 Mb/39 *VQs*) [36], *Arabidopsis* (125 Mb/34 *VQs*) [37], maize (2.3 Gb/61 *VQs*) [38] and *S. spontaneum* (3.36 Gb/78 *VQs*) [39], indicating that the number of *VQ* gene family members may not have an absolute correlation with genome size. While the AP85–441 was haploid and produced from the octoploid SES208, the number of *VQ* genes in octoploid *S. spontaneum* could be over 78 [5]. The gene structure analysis suggested that 70.5% (55/78) of *SsVQ* genes were found to be intronless. The same phenomenon has also been found in *Glycine max* (64; 85.33%) [17], *M. truncatula* (29; 90.63%) [40] and *G. raimondii* (31; 68.85%) [24]. Comparatively, these plants indicated that most *VQ* genes have lost introns during the long evolutionary period.

Based on the results of multiple sequence alignment, we found that there are four types of VQ domains for SsVQ proteins (LTG, FTG, VTG and ITG). Previous studies have shown that there are six types of motifs in *Arabidopsis* (LTG, FTG, VTG, YTG, LTS and LTD) [25], three types in grapevine (LTG, FTG, and VTG) [41], six types in maize (LTG, FTG, VTG, ITG, ATG and LTA) [19], four types in rice (LTG, FTG, VTG and ITG) [28] and seven types in tomato (LTG, FTG, VTG, LTS, LTA, YTG and HTG) [42]. Comparing the different plants of the conserved VQ domain, we found that LTG, FTG and VTG are the three most common domain types in plants. Even though maize and *S. spontaneum* are all

monocot plants, they have different VQ domain variations [19]. Therefore, the variations in conserved VQ motif may be different in various species.

The SsVQ proteins all contain the conserved VQ domain, except SsVQ66, SsVQ67 and SsVQ68, in which VH core amino acids replace the VQ core amino acids. The same phenomenon has also been found in rice and maize. The VQ core amino acids are also replaced by a VH core amino acid in rice and maize (OsVQ37, OsVQ39, ZmVQ15, ZmVQ28 and ZmVQ58) [19,28]. Interestingly, the VH core amino acids only showed in monocot plants, which may indicate the evolutionary process difference between monocot and dicots. Moreover, NtVQ54, in which isoleucine (IQ) replaces the canonical caline in the conserved domain (VQ), has demonstrated a similar phenomenon [22].

The major expansion ways of plant genome were segmental and tandem duplication events [34]. In this study, 24 gene pairs originated from segmental duplication and three gene pairs were derived from tandem duplication. Similarly, the phenomenon of low tandem and high segmental duplication proportion for gene families were also founded in *A. thaliana*, maize and Chinese cabbage [19,21,25]. In addition, gene duplication may lead to gene functional redundancy, and these duplicate genes could develop divergent patterns in gene expression for stable maintenance through subfunctionalization [43,44]. In our study, some paralogs showed different expression patterns, such as *SsVQ66/68*, *SsVQ24/30* and *SsVQ1/15*. Furthermore, there are numerous orthologous gene pairs in *SsVQ* genes. The substitution rates of Ka and Ks are the basis for analyzing the selection pressure in gene duplication events. The most Ka/Ks values of gene pairs were <1, suggesting that they had mainly undergone purifying selection.

Previous research has shown that the plant *VQ* gene family members have a different expression pattern, from growth and development to response to various external stresses, especially phytohormone treatment and cold stress [26]. Expression patterns for 78 *SsVQ* genes were conducted based on their tissue expression using RNA-seq data. The results showed that there were eight genes that had a relatively high expression in ten tissues, indicating that they may be related to the growth and development of plants [16,45]. Moreover, under drought and cold stress, the *SsVQs* expression levels were conducted using SRA data (Figure 8). The results showed that most of the *SsVQs* were up-regulated or down-regulated, which was consistent with the prediction that the promoter region contained a large number of LTR response element and drought response element. In addition, to verify the function of *SsVQs* in response to stress, we selected seven *SsVQs* for qRT-PCR analysis under four different stressors (SA, JA, ABA and cold). Most *SsVQ* genes were up-regulated with phytohormone treatment, and the results are consistent with *AtVQ* genes that can respond to SA treatment [25]. The expression of *SsVQ8* was robust and induced by three phytohormones of treatment, which is similar to *BrVQ23-2* and *BrVQ23-3* [21]. Under cold stress, the expression of *SsVQ57*, *SsVQ73* and *SsVQ76* were up-regulated, which consisted of SRA data. Together, these results suggested that *SsVQ* genes participate in the network of abiotic stress responses.

4. Materials and Methods

4.1. Identification and Classification of VQ Gene Family in Sugarcane

The genomic information of sugarcane was downloaded from the *S. spontaneum* AP85–441 genome (http://www.life.illinois.edu/ming/downloads/Spontaneumgenome/ accessed on 4 May 2021) [5]. The protein sequences of *VQ* genes of *A. thaliana* were downloaded from the NCBI (https://www.ncbi.nlm.nih.gov/ accessed on 4 May 2021). Firstly, the *A. thaliana* VQ proteins were used as query sequences to find SsVQ proteins of the *S. spontaneum* genome in the local BLAST (10^{-5}) and the VQ conserved domain (PF05678) was used as a query to explore the sugarcane proteins databases with the values (e-value) cut-off at 0.1. Next, all of the predicted VQ proteins were submitted to SMART (http://smart.emblheidelberg.de/ accessed on 6 May 2021) and NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi accessed on 10 May 2021) to confirm that they all contained the VQ motif. Lastly, physical parameters of

the VQ proteins, including protein length, isoelectric point (pI) and molecular weight were calculated in ExPASy (http://www.expasy.org/tools accessed on 10 May 2021). Subcellular localization was predicted using the WoLF PSORT (https://wolfpsort.hgc.jp/ accessed on 10 May 2021).

4.2. Phylogenetic Analysis and Multiple Sequence Alignment of SsVQ Proteins

To analyze the evolutionary relationship of the *VQ* gene families among *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays* and sugarcane, AtVQ, OsVQ, and ZmVQ proteins were downloaded from phytozomes (http://www.phytozome.org accessed on 15 May 2021). The neighbor-joining (NJ) phylogenetic tree was constructed using MEGA X software with 1000 bootstrap replications [46]. Genes were classified according to *Arabidopsis thaliana* and *Oryza sativa*. Multiple alignments of the SsVQ were conducted using Jalview software with default parameter settings based on full-length proteins [47].

4.3. Gene Structure, Conserved Domains and Motif Composition of SsVQ Genes

Conserved motifs of the VQ proteins were identified using MEME (v.5.3.3) (http://meme-suite.org/tools/meme accessed on 17 May 2021) with default settings [48]. Gene structure was investigated using GSDS 2.0 (http://gsds.cbi.pku.edu.cn/ accessed on 18 May 2021) [49]. Conserved domains of VQ proteins were identified using the NCBI CDD tool [50]. TBtools (v.1.0971) was used to integrate phylogenetic trees, conserved domains results, gene structure results and conserved motifs [51].

4.4. Promoter Analysis of SsVQ Genes

The 3000 bp sequences upstream of the *SsVQ* genes transcriptional start site were submitted to PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/ accessed on 20 May 2021) to identify the putative cis-elements [52].

4.5. Chromosomal Distribution and Duplication Analysis of SsVQ Genes

The chromosomal distribution in the identified *SsVQ* genes was obtained and visualized using TBtools software. The gene duplication events were conducted by using MCScanX [53]. The duplication events were also detected for the *SsVQ* genes. Two proteins that have more than 40% similarity and were separated by four or fewer gene loci were identified as tandem duplication, and others were identified as segmental duplications. [54]. Non-synonymous (ka) and synonymous (ks) substitutions of each duplicated *SsVQ* genes were calculated using KaKs_Calculator 2.0 [55].

4.6. Prediction of the Protein-Protein Interaction Network

In total, 78 VQ proteins were submitted to STRING (version 11.5, https://cn.string-db.org/ accessed on 22 May 2021) as queries, and *Zea mays* was chosen as the reference genome for blasting with default settings [56,57]. Then, the network was constructed by the best matched homologs in *Zea mays*.

4.7. Expression Analysis of SsVQs Based on RNA-Seq

First, different sugarcane tissue expression data, including leaves and stems at 35 days (seedling stage), 9 months (early maturity stage) and 12 months (maturity stage), was download from *Saccharum* Genome database (http://sugarcane.zhangjisenlab.cn/sgd/html/index.html accessed on 6 June 2021). In addition, under drought and cold treatments, the public data (PRJNA590595 and PRJNA636260) were download from the SRA database (https://www.ncbi.nlm.nih.gov/sra/ accessed on 6 May 2021). Fastq [58] and hisat2 [59] tools were applied to improve the sequence quality and map sequence data on the reference genome (*S. spontaneum* AP85–441), respectively. The featurCounts in Subread package and trimmed mean of M-values (TMMs) were conducted for count read and normalization [60,61]. Finally, all the expression values (FPKMs or TMMs) were used to create heat maps and cluster analysis through TBtools [51].

4.8. Plant Materials, Phytohormone and Cold Treatments

The sugarcane cultivar ROC22 (*Saccharum* spp. hybrid) was provided by Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture, Fujian Agriculture and Forestry University (Fuzhou 350002, China). Uniform 20-day-old seedlings with 3–4 fully unfolded leaves were cultured in 1/4 Hoagland nutrient solution for one week, and then these plantlets were divided into four groups. The three groups were sprayed with ABA (100 μmol/L), SA (5 mmol/L) and MeJA (100 μmol/L), respectively (Sangon Biotech, Shanghai, China). The leaves were harvested at 0, 1, 6 and 12 h. Meanwhile, one group was treated with low temperature (4 °C). Then the leaves were harvested at 0, 1, 6 and 24 h. All samples were subject to flash freezing in liquid nitrogen immediately after collection and subsequently stored in the refrigerator at −80 °C until RNA extraction.

4.9. Quantitative Real-Time PCR (qRT-PCR) Analysis

Total RNA was extracted from each sample using RNAiso Plus (Takara, Dalian, China) according to the manufacturer's instructions. The cDNA synthesis was carried out with approximately 2 μg RNA using PrimeScript RT reagent Kit with gDNA Eraser (Takara, Dalian, China). The 15 μL reaction system of qRT-PCR contains 7.5 μL SYBR Green Master Mix, 0.6 μL forward primers, 0.6 μL reverse primers, 1 μL cDNA template and 5.3 μL sterile distilled water. The qRT-PCR procedure was 95 °C for 3 min, 45 cycles of 94 °C for 15 s and 60 °C for 30 s. After that, the melting curves were analyzed. For each sample, three technical replicates were conducted to calculate the averaged Ct values. The 25S genes were used as the internal control. The $2^{-\Delta\Delta Ct}$ method was used to calculate the qRT-PCR data. Primers used for this study are listed in Table S7.

5. Conclusions

This study comprehensively analyzed the *VQ* gene family of sugarcane. A total of 78 *SsVQs* were identified and classified in detail. Five types of VQ specificity domain were found by multiple sequence alignment, especially the FxxxVHQxVTG type, which were identified in gene *SsVQ66*, *SsVQ67* and *SsVQ68.* Gene duplication analysis showed that *VQ* genes had expanded mainly through whole-genome segmental duplication in *Saccharum spontaneum*. The promoter regions of the *SsVQ* genes contained an extremely large number of the LTR response element, drought response element and phytohormone response element, to respond to various stresses. The RNA-seq analysis showed that 23 *VQ* genes were expressed in all 10 tissues, indicating that they may be related to the growth and development of sugarcane. SRA and qRT-PCR analysis showed that genes such as *SsVQ57*, *SsVQ73* and *SsVQ76* were selected as candidate genes for agricultural purposes, which obviously respond to cold, drought and phytohormone. These results provide valuable information to understand the biological role of the *SsVQ* genes and candidate genes for molecular-assisted breeding of sugarcane.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23116334/s1.

Author Contributions: Conceptualization, Y.L., L.X. and D.Z. (Dinggang Zhou); methodology, Y.Q.; software, Y.L., X.L., D.Y., Y.J. and Z.Y.; validation, J.W., D.Z. (Dawei Zhang) and L.L.; writing—original draft preparation, Y.L., X.L. and D.Z. (Dinggang Zhou); writing—review and editing, H.L., N.H., M.Y., Y.Q. and D.Z. (Dinggang Zhou); project administration, L.X., Y.Q. and D.Z. (Dinggang Zhou). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by "National Natural Science Foundation of China, grant number U19A2029 and 31701491", "the open funding of National Engineering Research Center for Sugarcane, Fujian Agriculture and Forestry University, grant number 2017.6.2".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated for this study can be found in SRA, the accession number: PRJNA590595 and PRJNA636260.

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

References

1. Zhou, D.; Liu, X.; Gao, S.; Guo, J.; Su, Y.; Ling, H.; Wang, C.; Li, Z.; Xu, L.; Que, Y. Foreign *cry1Ac* gene integration and endogenous borer stress-related genes synergistically improve insect resistance in sugarcane. *BMC Plant Biol.* **2018**, *18*, 342. [CrossRef] [PubMed]
2. Xu, F.; He, L.; Gao, S.; Su, Y.; Li, F.; Xu, L. Comparative analysis of two sugarcane ancestors *Saccharum officinarum* and *S. spontaneum* based on complete chloroplast genome sequences and photosynthetic ability in cold stress. *Int. J. Mol. Sci.* **2019**, *20*, 3828. [CrossRef] [PubMed]
3. Liu, X.; Xu, L.; Que, Y.; Ling, H.; Yin, Z.; Liu, Y.; Zhou, D. The complete mitochondrial genome and phylogenetic analysis of sugarcane (*Saccharum* spp. hybrids) line 15a-53. *Mitochondrial DNA B Resour.* **2020**, *5*, 3389–3391. [CrossRef] [PubMed]
4. Li, Z.; Hua, X.T.; Zhong, W.M.; Yuan, Y.; Wang, Y.J.; Wang, Z.C.; Ming, R.; Zhang, J.S. Genome-wide identification and expression profile analysis of WRKY family genes in the autopolyploid *Saccharum spontaneum*. *Plant Cell Physiol.* **2020**, *61*, 616–630. [CrossRef]
5. Zhang, J.; Zhang, X.; Tang, H.; Zhang, Q.; Hua, X.; Ma, X.; Zhu, F.; Jones, T.; Zhu, X.; Bowers, J.; et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat. Genet.* **2018**, *50*, 1565–1573. [CrossRef] [PubMed]
6. Lembke, C.G.; Nishiyama, M.Y., Jr.; Sato, P.M.; de Andrade, R.F.; Souza, G.M. Identification of sense and antisense transcripts regulated by drought in sugarcane. *Plant Mol. Biol.* **2012**, *79*, 461–477. [CrossRef]
7. Huang, X.; Song, X.P.; Chen, R.F.; Zhang, B.Q.; Li, C.N.; Liang, Y.S.; Qiu, L.H.; Fan, Y.G.; Zhou, Z.F.; Zhou, H.W.; et al. Genome-wide analysis of the DREB subfamily in *Saccharum spontaneum* reveals their functional divergence during cold and drought stresses. *Front. Genet.* **2020**, *10*, 1326. [CrossRef]
8. Rice, J.L.; Hoy, J.W.; Grisham, M.P. Sugarcane mosaic distribution, incidence, increase, and spatial pattern in Louisiana. *Plant Dis.* **2019**, *103*, 2051–2056. [CrossRef]
9. Dessoky, E.S.; Ismail, R.M.; Elarabi, N.I.; Abdelhadi, A.A.; Abdallah, N.A. Improvement of sugarcane for borer resistance using *Agrobacterium* mediated transformation of *cry1Ac* gene. *GM Crops Food* **2021**, *12*, 47–56. [CrossRef]
10. Basnayake, J.; Jackson, P.A.; Inman-Bamber, N.G.; Lakshmanan, P. Sugarcane for water-limited environments. Genetic variation in cane yield and sugar content in response to water stress. *J. Exp. Bot.* **2012**, *63*, 6023–6033. [CrossRef]
11. Viswanathan, R.; Balamuralikrishnan, M. Impact of mosaic infection on growth and yield of sugarcane. *Sugar Tech* **2005**, *7*, 61–65. [CrossRef]
12. Lu, G.; Wang, Z.; Xu, F.; Pan, Y.-B.; Grisham, M.P.; Xu, L. Sugarcane mosaic disease: Characteristics, identification and control. *Microorganisms* **2021**, *9*, 1984. [CrossRef] [PubMed]
13. Wang, W.Z.; Yang, B.P.; Feng, X.Y.; Cao, Z.Y.; Feng, C.L.; Wang, J.G.; Xiong, G.R.; Shen, L.B.; Zeng, J.; Zhao, T.T.; et al. Development and characterization of transgenic sugarcane with insect resistance and herbicide tolerance. *Front. Plant Sci.* **2017**, *8*, 1535. [CrossRef] [PubMed]
14. Qamar, Z.; Nasir, I.A.; Abouhaidar, M.G.; Hefferon, K.L.; Rao, A.Q.; Latif, A.; Ali, Q.; Anwar, S.; Rashid, B.; Shahid, A.A. Novel approaches to circumvent the devastating effects of pests on sugarcane. *Sci. Rep.* **2021**, *11*, 12428. [CrossRef]
15. Jones, J.D.; Dangl, J.L. The plant immune system. *Nature* **2006**, *444*, 323–329. [CrossRef] [PubMed]
16. Cai, H.; Zhang, M.; Liu, Y.; He, Q.; Chai, M.; Liu, L.; Chen, F.; Huang, Y.; Yan, M.; Zhao, H.; et al. Genome-wide classification and evolutionary and functional analyses of the *VQ* family. *Trop. Plant Biol.* **2019**, *12*, 117–131. [CrossRef]
17. Wang, Y.; Jiang, Z.; Li, Z.; Zhao, Y.; Tan, W.; Liu, Z.; Cui, S.; Yu, X.; Ma, J.; Wang, G.; et al. Genome-wide identification and expression analysis of the *VQ* gene family in soybean (*Glycine max*). *PeerJ* **2019**, *7*, e7509. [CrossRef] [PubMed]
18. Lai, Z.; Li, Y.; Wang, F.; Cheng, Y.; Fan, B.; Yu, J.Q.; Chen, Z. Arabidopsis sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. *Plant Cell* **2011**, *23*, 3824–3841. [CrossRef]
19. Song, W.; Zhao, H.; Zhang, X.; Lei, L.; Lai, J. Genome-wide identification of VQ motif-containing proteins and their expression profiles under abiotic stresses in maize. *Front. Plant Sci.* **2015**, *6*, 1177. [CrossRef] [PubMed]
20. Ye, Y.J.; Xiao, Y.Y.; Han, Y.C.; Shan, W.; Fan, Z.Q.; Xu, Q.G.; Kuang, J.F.; Lu, W.J.; Lakshmanan, P.; Chen, J.Y. Banana fruit VQ motif-containing protein5 represses cold-responsive transcription factor MaWRKY26 involved in the regulation of JA biosynthetic genes. *Sci. Rep.* **2016**, *6*, 23632. [CrossRef]
21. Zhang, G.; Wang, F.; Li, J.; Ding, Q.; Zhang, Y.; Li, H.; Zhang, J.; Gao, J. Genome-wide identification and analysis of the VQ motif-containing protein family in Chinese cabbage (*Brassica rapa* L. ssp. *Pekinensis*). *Int. J. Mol. Sci.* **2015**, *16*, 28683–28704. [CrossRef] [PubMed]
22. Liu, C.; Liu, H.; Zhou, C.; Timko, M.P. Genome-wide identification of the VQ protein gene family of tobacco (*Nicotiana tabacum* L.) and analysis of its expression in response to phytohormones and abiotic and biotic stresses. *Genes* **2020**, *11*, 284. [CrossRef]
23. Yan, H.; Wang, Y.; Hu, B.; Qiu, Z.; Zeng, B.; Fan, C. Genome-wide characterization, evolution, and expression profiling of VQ gene family in response to phytohormone treatments and abiotic stress in *Eucalyptus grandis*. *Int. J. Mol. Sci.* **2019**, *20*, 1765. [CrossRef] [PubMed]

24. Chen, P.; Wei, F.; Cheng, S.; Ma, L.; Wang, H.; Zhang, M.; Mao, G.; Lu, J.; Hao, P.; Ahmad, A.; et al. A comprehensive analysis of cotton *VQ* gene superfamily reveals their potential and extensive roles in regulating cotton abiotic stress. *BMC Genom.* **2020**, *21*, 795. [CrossRef] [PubMed]
25. Cheng, Y.; Zhou, Y.; Yang, Y.; Chi, Y.J.; Zhou, J.; Chen, J.Y.; Wang, F.; Fan, B.; Shi, K.; Zhou, Y.H.; et al. Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors. *Plant Physiol.* **2012**, *159*, 810–825. [CrossRef] [PubMed]
26. Jing, Y.; Lin, R. The VQ motif-containing protein family of plant-specific transcriptional regulators. *Plant Physiol.* **2015**, *169*, 371–378. [CrossRef] [PubMed]
27. Dong, Q.; Zhao, S.; Duan, D.; Tian, Y.; Wang, Y.; Mao, K.; Zhou, Z.; Ma, F. Structural and functional analyses of genes encoding VQ proteins in apple. *Plant Sci.* **2018**, *272*, 208–219. [CrossRef] [PubMed]
28. Kim, D.Y.; Kwon, S.I.; Choi, C.; Lee, H.; Ahn, I.; Park, S.R.; Bae, S.C.; Lee, S.C.; Hwang, D.J. Expression analysis of rice VQ genes in response to biotic and abiotic stresses. *Gene* **2013**, *529*, 208–214. [CrossRef]
29. Jiang, S.Y.; Sevugan, M.; Ramachandran, S. Valine-glutamine (VQ) motif coding genes are ancient and non-plant-specific with comprehensive expression regulation by various biotic and abiotic stresses. *BMC Genom.* **2018**, *19*, 342. [CrossRef]
30. Wang, A.; Garcia, D.; Zhang, H.; Feng, K.; Chaudhury, A.; Berger, F.; Peacock, W.J.; Dennis, E.S.; Luo, M. The VQ motif protein IKU1 regulates endosperm growth and seed size in *Arabidopsis*. *Plant J.* **2010**, *63*, 670–679. [CrossRef] [PubMed]
31. Hu, P.; Zhou, W.; Cheng, Z.; Fan, M.; Wang, L.; Xie, D. JAV1 controls jasmonate-regulated plant defense. *Mol. Cell* **2013**, *50*, 504–515. [CrossRef]
32. Perruc, E.; Charpenteau, M.; Ramirez, B.C.; Jauneau, A.; Galaud, J.P.; Ranjeva, R.; Ranty, B. A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. *Plant J.* **2004**, *38*, 410–420. [CrossRef] [PubMed]
33. Xixu, P.; Ting, X.; Jiao, M.; Zong, T.; Dinggang, Z.; Xinke, T.; Haihua, W. Differential expression of rice valine-qlutamine gene family in response to nitric oxide and regulatory circuit of *OsVQ7* and *OsWRKY24*. *Rice Sci.* **2020**, *27*, 10–20. [CrossRef]
34. Kaltenegger, E.; Leng, S.; Heyl, A. The effects of repeated whole genome duplication events on the evolution of cytokinin signaling pathway. *BMC Evol. Biol.* **2018**, *18*, 76. [CrossRef] [PubMed]
35. Liu, X.; Yin, Z.; Liu, Y.; Li, Z.; Yan, M.; Que, Y.; Xu, L.; Zhou, D. The complete mitochondrial genome of sugarcane (*Saccharum* spp.) variety FN15. *Mitochondrial DNA Part B* **2020**, *5*, 2163–2165. [CrossRef] [PubMed]
36. Yu, J.; Ni, P.; Wong, G.K. Comparing the whole-genome-shotgun and map-based sequences of the rice genome. *Trends Plant Sci.* **2006**, *11*, 387–391. [CrossRef]
37. Arabidopsis Genome, I. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **2000**, *408*, 796–815. [CrossRef] [PubMed]
38. Xiao, Y.; Liu, H.; Wu, L.; Warburton, M.; Yan, J. Genome-wide association studies in maize: Praise and Stargaze. *Mol. Plant* **2017**, *10*, 359–374. [CrossRef]
39. Garsmeur, O.; Droc, G.; Antonise, R.; Grimwood, J.; Potier, B.; Aitken, K.; Jenkins, J.; Martin, G.; Charron, C.; Hervouet, C.; et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* **2018**, *9*, 2638. [CrossRef] [PubMed]
40. Ling, L.; Qu, Y.; Zhu, J.; Wang, D.; Guo, C. Genome-wide identification and expression analysis of the VQ gene family in *Cicer arietinum* and *Medicago truncatula*. *PeerJ* **2020**, *8*, e8471. [CrossRef]
41. Wang, M.; Vannozzi, A.; Wang, G.; Zhong, Y.; Corso, M.; Cavallini, E.; Cheng, Z.M. A comprehensive survey of the grapevine *VQ* gene family and its transcriptional correlation with WRKY proteins. *Front. Plant Sci.* **2015**, *6*, 417. [CrossRef] [PubMed]
42. Ding, H.; Yuan, G.; Mo, S.; Qian, Y.; Wu, Y.; Chen, Q.; Xu, X.; Wu, X.; Ge, C. Genome-wide analysis of the plant-specific VQ motif-containing proteins in tomato (*Solanum lycopersicum*) and characterization of SlVQ6 in thermotolerance. *Plant Physiol. Biochem.* **2019**, *143*, 29–39. [CrossRef] [PubMed]
43. Zhang, J. Evolution by gene duplication: An update. *Trends Ecol. Evol.* **2003**, *18*, 292–298. [CrossRef]
44. Krishnamurthy, P.; Hong, J.K.; Kim, J.A.; Jeong, M.J.; Lee, Y.H.; Lee, S.I. Genome-wide analysis of the expansin gene superfamily reveals *Brassica rapa*-specific evolutionary dynamics upon whole genome triplication. *Mol. Genet. Genom.* **2015**, *290*, 521–530. [CrossRef]
45. Wang, X.; Zhang, H.; Sun, G.; Jin, Y.; Qiu, L. Identification of active VQ motif-containing genes and the expression patterns under low nitrogen treatment in soybean. *Gene* **2014**, *543*, 237–243. [CrossRef]
46. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
47. Clamp, M.; Cuff, J.; Searle, S.M.; Barton, G.J. The Jalview Java alignment editor. *Bioinformatics* **2004**, *20*, 426–427. [CrossRef]
48. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [CrossRef] [PubMed]
49. Hu, B.; Jin, J.; Guo, A.-Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296–1297. [CrossRef] [PubMed]
50. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [CrossRef]

51. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef] [PubMed]
52. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [CrossRef] [PubMed]
53. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H.; et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef]
54. Li, P.T.; Chai, Z.; Lin, P.P.; Huang, C.H.; Huang, G.Q.; Xu, L.N.; Deng, Z.H.; Zhang, M.Q.; Zhang, Y.; Zhao, X.W. Genome-wide identification and expression analysis of AP2/ERF transcription factors in sugarcane (*Saccharum spontaneum* L.). *BMC Genom.* **2020**, *21*, 685. [CrossRef] [PubMed]
55. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [CrossRef]
56. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **2021**, *49*, D605–D612. [CrossRef] [PubMed]
57. Su, W.; Zhang, C.; Wang, D.; Ren, Y.; Zhang, J.; Zang, S.; Zou, W.; Su, Y.; You, C.; Xu, L.; et al. A comprehensive survey of the aldehyde dehydrogenase gene superfamily in *Saccharum* and the role of ScALDH2B-1 in the stress response. *Environ. Exp. Bot.* **2022**, *194*, 104725. [CrossRef]
58. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef]
59. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]
60. Liao, Y.; Smyth, G.K.; Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **2019**, *47*, e47. [CrossRef] [PubMed]
61. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **2010**, *11*, R25. [CrossRef] [PubMed]