Genome-wide (ChIP-seq) Identification of Target Genes Regulated by WRKY33 During Submergence Stress in *Arabidopsis*

Zhang Junlin
Sichuan University - Wangjiang Campus: Sichuan University
Liu bao
Sichuan University - Wangjiang Campus: Sichuan University
Song Yan
Sichuan University - Wangjiang Campus: Sichuan University
Chen yang
Sichuan University - Wangjiang Campus: Sichuan University
Fu jiao
Sichuan University - Wangjiang Campus: Sichuan University
Liu Jianquan
Sichuan University - Wangjiang Campus: Sichuan University
Ma Tao
Sichuan University - Wangjiang Campus: Sichuan University
Xi Zhenxiang
Sichuan University - Wangjiang Campus: Sichuan University
Liu Huanhuan (liuhuanhuan85@163.com)
Sichuan University - Wangjiang Campus: Sichuan University https://orcid.org/0000-0003-1790-3073

Research article

Keywords: WRKY33, submergence treatment, hypoxia, ChIP-seq, Arabidopsis

Posted Date: December 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-125855/v1

Abstract

Background: Hypoxia induced by flooding causes significant losses to crop production almost every year. However, the molecular network of hypoxia signaling pathway is still poorly understood. According to previous studies, transgenic plants overexpressing the *WRKY33* gene showed enhanced resistance to submergence stress. Thus, the transcription factor WRKY33 may regulate a series of target genes in response to submergence. Here, to determine the putative downstream targets of WRKY33 at a genome-wide scale in *Arabidopsis thaliana*, we performed the chromatin immunoprecipitation sequencing (ChIP-seq) using *35S:FLAG-WRKY33* overexpression transgenic lines (*WRKY33-OE*) upon 24 hours of submergence treatment.

Results: Using ChIP-seq data, we identified a total of 104 WRKY33-binding genes under submergence treatment (WRKY33BGHs). Most of the WRKY33BGHs are involved in the oxidation-reduction process, programmed cell death in response to reactive oxygen species, lipid biosynthesis process, and other processes related to stress responses. Moreover, the major motif identified in the WRKY33BGHs promoters is a new *cis*-element, TCTCTC (we named it as “TC box”). This *cis*-element is different from the previously known W box for WRKY33. Further qPCR experiments verified that genes carrying this motif in their promoters could be regulated by WRKY33 upon submergence treatment.

Conclusions: Our study has identified a new putative binding motif of WRKY33 and recovered numerous previously unknown target genes of WRKY33 during submergence stress. The *WRKY33* gene positively participates in flooding response probably by transcriptional regulation of the downstream hypoxia-related target genes via a “TC box”.

Background

Large areas of cropland in the world are subject to seasonal flooding, which causes significant losses to crop production almost every year. The diffusion of oxygen in water is 10,000 times slower than that in air [1], drastically reducing the supply of oxygen to the plants. Morphological adaptations of plants to low-oxygen stress include the formation of adventitious roots, as well as the development of cortical air spaces in roots that promote air transport [2]. Meanwhile, the induction of fermentation pathway enzymes has been established as an important metabolic adaptation to anaerobiosis [3, 4]. Over the last decade, it has become increasingly evident that the N-degron pathway plays a well-characterized role in the response to hypoxia, e.g., after flooding and plant submergence [5, 6]. In addition, a variety of transcription factors (TFs) have been reported to regulate gene expression that promotes adaptive responses to the environmental and physiological stress [7], including the *Dof* (DNA-binding with one finger) gene family [8], the MADS-box gene family [9], and the *WRKY* gene family.

The WRKY TF family, found exclusively in green plants, is characterized by the highly conserved amino acid sequence WRKYGQK at the N-terminus and the zinc-finger structure at the C-terminus [10]. Numerous studies have demonstrated that WRKY TFs are involved in regulation of various processes, such as seed

germination, leaf senescence, and the responses to biotic and abiotic stresses [11, 12]. In particular, one member of the WRKY TF family, WRKY33, has been shown to regulate plant defense responses to a variety of stresses [13, 14]. For example, previous studies have documented that overexpression of the *WRKY33* gene enhances the resistance to oxidative stress [15] and promotes pathogen defense [16]. In addition, a recent study has revealed that WRKY33 positively regulates submergence-induced hypoxia in *Arabidopsis thaliana* (L.) Heynh. and directly acts on the *RAP2.2* gene, which encodes a TF belonging to the group VII Ethylene Response Factors (ERFVIIs) [17]. Despite these studies, the downstream regulatory network governed by WRKY33 is still poorly understood. Here, using chromatin immunoprecipitation followed by high-throughput sequencing (i.e., ChIP-seq), we aim to identify all WRKY33-targeted genes in response to submergence.

Results

Verification of the function and phenotype of 35S:FLAG-WRKY33 transgenic Arabidopsis in submergence response

A previous study showed that *WRKY33* was induced by hypoxia stress in roots of *Arabidopsis* [4]. Recently, *WRKY33* was reported to positively regulate submergence response via interacting with WRKY12 to directly upregulate *RAP2.2* in *Arabidopsis* [17]. To further identify other WRKY33 targeted genes during submergence response at a genome-wide scale, we aimed to conduct the chromatin immunoprecipitation followed by high throughput sequencing (ChIP-Seq) using *35S:FLAG-WRKY33* overexpression transgenic plants (*WRKY33-OE*) upon 24 hours' submergence treatment. Before the ChIP experiment, we obtained the *WRKY33OE* transgenic plants (Supplemental Fig. 1) in Col background and examined its submergence tolerance ability to make sure that the plants were workable. The phenotypic assay showed that *WRKY33OE* plants were more tolerant to submergence treatment compared to Col (Supplemental Fig. 2A). Survival rates and dry weights of Col, *WRKY33OE-1* and *WRKY33OE-2* plants were also consistent with their phenotypic assays (Supplemental Fig. 2B-C). Malondialdehyde (MDA) contents (Supplemental Fig. 2D) were also evaluated among Col, *WRKY33OE-1* and *WRKY33OE-2* plants and the results also supported that overexpression of *FLAG-WRKY33* enhanced the submergence tolerance ability in *Arabidopsis*. These results together showed that overexpression of *FLAG-WRKY33* could enhance the submergence tolerance ability compared to wild type and the *WRKY33-OE* transgenic plants could be used for the following ChIP experiment.

Analysis of the ChIP-Seq peaks

Having confirmed that the *WRKY33OE* transgenic plants had the enhanced submergence resistance ability, we then performed the ChIP experiment firstly by using the samples (2 g pooled leaf materials) of 14-day-old seedlings of *WRKY33OE1* and *WRKY33OE2* plants. The average size of the input fragments and the anti-FLAG chip libraries were approximately 100–400 bp. The immunoprecipitated DNA fragments were then sent to the BGI (Shenzhen, China) company for further sequencing. The input library had 25.4 million reads and the FLAG Ab ChIP library had 24.6 million reads. More than 95% of the reads

were mapped to the *Arabidopsis* genome. The MACS2 program (Analysis based on ChIP-Seq models) [18] was used to identify the enriched regions using a false discovery cutoff of 0.05. The location of the enriched peaks in the *Arabidopsis* genome is shown in the supplemental table 1 (Additional file 1). Of the 393 enriched regions, 24% of the peaks were in genetic regions (from 2 kb upstream of the start of transcription to 2 kb downstream of the stop codon, including the coding region). Of the peaks that were in the genetic regions, 22% located only in the promoter regions, 48% in the promoter and exons or introns regions, only 26% in exons and introns (Fig. 1). After calling peak, we aimed to examine the peak locations among the whole genome. We then used the covplot function in ChIPseeker (an R package for ChIP peak Annotation, Comparison and Visualization) to calculate the coverage of peak regions over the chromosomes. We generated a figure for visualization (Fig. 2A). Since some annotations overlapped, we then viewed the complete annotations with overlap through the vennpie function in ChIPseeker (Fig. 2B). Table 1 lists the genes related to the peaks in the gene region. These peaks are enriched by more than 5-fold and all have known putative functions.

Table 1
List of genes and their putative function

Gene Name	Putative Function	
AT1G21650	Preprotein translocase SecA family protein	8.0
AT1G64628	conserved peptide upstream open reading frame 57	5.2
AT2G01008	maternal effect embryo arrest protein	8.1
AT2G15540	non-LTR retrotransposon family	5.2
AT2G18220	Noc2p family	5.2
AT2G29350	senescence-associated gene 13	7.7
AT2G31040	Encodes an integral thylakoid protein that facilitates assembly of the membranous part of the chloroplast ATPase	9.6
AT2G47090	zinc ion binding/nucleic acid binding protein	12.4
AT3G10810	zinc finger (C3HC4-type RING finger) family protein	6.4
AT3G11280	Duplicated homeodomain-like superfamily protein	5.9
AT3G11900	aromatic and neutral transporter 1	7.1
AT3G12120	fatty acid desaturase 2	7.8
AT3G22160	JAV1 is a repressor of jasmonate-mediated defense responses	11.3
AT3G22170	far-red elongated hypocotyls 3	9.5
AT3G27503	Encodes a member of a family of small, secreted, cysteine rich proteins with sequence similarity to SCR	10.6
AT3G30250	transposable element gene	8.0
AT3G33058	gypsy-like retrotransposon family	15.7
AT3G41768	rRNA	10.2
AT3G41979	5.8SrRNA	6.8
AT3G42130	glycine-rich protein	6.2
AT3G45755	transposable element gene	6.1
AT3G52140	tetratricopeptide repeat (TPR)-containing protein	6.2
AT4G10030	Alpha/beta hydrolase domain containing protein involved in lipid biosynthesis	5.3
AT4G20360	Nuclear transcribed, plastid localized EF-Tu translation elongation factor	5.2

The genes listed in this table are limited to those associated with peaks that were enriched greater than 5-fold and have been classified with a known function

Gene Name	Putative Function	
AT4G32700	helicases;ATP-dependent helicases;nucleic acid binding;ATP binding;DNA-directed DNA polymerases;DNA binding	5.3
AT4G32810	carotenoid cleavage dioxygenase 8	5.2
AT4G34035	pre-tRNA tRNA-Arg	9.6
AT4G34040	RING/U-box superfamily protein	7.9
AT4G35090	catalase 2	5.2
AT4G39672	pre-tRNA	6.1
AT5G17420	Encodes a xylem-specific cellulose synthase that is phosphorylated on one or more serine residues	30.1
AT5G17730	P-loop containing nucleoside triphosphate hydrolases superfamily protein	8.0
AT5G18650	CHY-type/CTCHY-type/RING-type Zinc finger protein	8.5
AT5G37960	GroES-like family protein	5.4
AT5G40690	histone-lysine N-methyltransferase trithorax-like protein	6.1
AT5G61710	cotton fiber protein	5.3
AT5G61710	cotton fiber protein	5.3
The genes listed in this table are limited to those associated with peaks that were enriched greater than 5-fold and have been classified with a known function		

Motif analysis of WKRY33 TF targeted genes

We analyzed all the promoter-located peak sequences from the ChIP-Seq using MEME-ChIP [19, 20] to identify the enriched motif, and detected the two types of motifs (Fig. 3A). The most significantly enriched MEME motif is “TCTCTCTC” (E-value of 6.3e-005) which is different from the “W box” bound by WRKY33 TF reported previously. We then named it as “TC box” (Fig. 3B). The next most significant motif is AAAAWAAA (E-value of 3.1e + 002) (Fig. 3C). WRKY proteins can repress or activate the expression of downstream genes via binding to the W-box (TGACC (A/T)) in promoter of its target genes upon pathogen defense [18]. The identified “TC box” motif may responsible for the activation or repression of submergence-related target genes which still needs further verifications.

Gene ontology analysis to identify biological and functional enriched categories

Gene Ontology (GO) analyses using the Enrich GO [21] revealed 61 GO categories belonging to the Biological Process (BP) ontology, which were determined to be significantly over-represented in the ChIP-Seq sample relative to the *Arabidopsis* genome (fisher < 0.01, Additional file 2). The top 10 significantly

enriched GO biological processes of WRKY33BGHs were shown in Fig. 4A. The results of the top 20 extremely significant enrichments (Fig. 4B) suggest that the gene ontology related to the submergence response includes the oxidation-reduction process, programmed cell death in response to reactive oxygen species and lipid biosynthesis process. Additional biological processes including cellular response to auxin stimulus, response to hydrogen peroxide were also identified when using a fisher greater than 0.01 and less than 0.05 (Additional file 2). Plant phytohormones, such as auxin, may also participate in the submergence response process as suggested by our Gene Ontology (GO) analysis, which still needs further experimental validation.

Expression analysis of genes contain the "TC box" in Col and WRKY33OE plants after submergence treatment

Genes contain the "TC box" may be the direct targets of WRKY33 during submergence response. To further validate this hypothesis, we selected four genes that contain the "TC box" and performed the QPCR test experiment. The results showed the expression levels of these four genes were all regulated by WRKY33 transcription factor. *At2G35736* gene was downregulated by WRKY33 while the other three genes *At1G66810*, *At2G47090*, and *At3g12120* were upregulated by WRKY33 (Fig. 5). These results support that these four genes targeted by WRKY33 may participate in submergence response via the "TC box". However, further experimental validations through more detailed experiments are needed in the future.

Discussion

Flooding stress, one of the most important abiotic stresses, has attracted the attention of scientists over the world [22]. Many studies have revealed the molecular mechanisms of plants in response to flooding [22], and WRKY33 has been reported as a positive regulator in responses to various abiotic stresses [23, 24]. By phenotypic analysis, we found that plants overexpressing *FLAG-WRKY33* did enhance resistance to hypoxia stress induced by submergence treatment compared with Col (Supplemental Fig. 2). We further performed the chromatin immunoprecipitation sequencing (ChIP-Seq) using *35S:FLAG-WRKY33* overexpressing transgenic lines (*WRKY33-OE*) under submergence treatment, to identify the WRKY33 TF target genes during submergence at a genome-wide scale. By ChIP-Seq analyses, we identified 104 WRKY33-binding genes under submergence treatment (WRKY33BGHs) and gene enrichment analysis showed that these genes participated in oxidoreductase reactions, lipid biosynthetic process and other functions. Most of these identified genes are reported for the first time for submergence stress. The major motif that we identified in the WRKY33BGHs promoters is the "TC box" cis-element. This candidate motif for WRKY33 TF may regulate genes expression during submergence stress. Our further functional analyses of all identified genes suggest that WRKY33BGHs may protect cells from oxidative stress and other processes to improve the tolerance ability upon submergence stress.

The identified "TC box" cis-element is a new motif different from the known "W box" element for WRKY33 and may be specific to regulating the target genes during submergence stress. WRKY33 can regulate

RAP2.2 expression via the W box element only during the submergence response [17]. Interestingly, there also is a “TC box” sequence “TCTCTC” in the promoter region (-1,875 bp) of *RAP2.2*. Previous studies have shown that the TFs have different binding abilities towards different cis-elements upon different conditions. For example, IPA1 was reported to bind to the “GTAC” element in the promoter of *DEP1* in the normal condition while bind to the “TGGGCC” element in the promoter of *WRKY45* upon pathogen infection [25]. This switch is mediated by the phosphorylation in IPA1 protein. Submergence treatment might also can induce the phosphorylation of WRKY33 like IPA1 upon pathogen infection [17, 22]. In addition, this TF may also have different binding abilities towards “W box” or “TC box” elements between normal growth and submergence treatment conditions like IPA1. Such a difference in binding ability may be mediated by the protein post-transcriptional modifications of WRKY33.

In this study, we obtained a more comprehensive understanding of the submergence stress response mediated by WRKY33. The ChIP-seq candidate genes regulated by WRKY33 provide a more comprehensive understanding of the molecular basis of plant submergence response. These genes can be further manipulated to improve stress tolerances when their functions and regulation pathways are well clarified. In addition, functions of these genes induced by low-oxygen stress seem to be overlapped with those by other biotic or abiotic stress responses [26]. Therefore, the hypoxic response may be more complex than previously expected and involve more adaptations except for shift of energy metabolism [27].

Conclusion

WRKY33 positively participates in flooding response probably by transcriptional directly activating its downstream hypoxia-related target genes via a “TC box” element.

Methods

Arabidopsis growing conditions and submergence treatment

Briefly, cDNA was prepared from 4-week rosette leaves of *Arabidopsis* and was diluted to 50 times. The diluted cDNA was then used as a template to amplify the *WRKY33*, which was inserted into a vector tagged by FLAG tag, under the control of the 35S promoter. The construct was transformed into *Agrobacterium* strain GV3101 [28], which was used to transform *Arabidopsis* using the floral dip method and identified by hygromycin screening followed by qRT-PCR analysis of their expression levels. All materials were grown at 22°C in a 16-hour light/8-hour dark cycle. Seeds were germinated on 1/2 MS medium (pH = 5.85) for seven days and then transplanted into soil.

For submergence treatments, 4-week-old plants were submerged 10 cm below the surface of the water in darkness for 65 hours. All submergence treatments started at 9:00 a.m. Twelve Col and *WRKY33OE* plants were used for submergence treatment every time. The total experiments were repeated three times.

For ChIP-sequencing, 4-week-old *35S:FLAG-WRKY33* transgenic plants were submerged 10 cm below the surface of the water in darkness for 24 hours. Then rosette leaves were collected for ChIP experiments. All submergence treatments started at 9:00 a.m.

Malondialdehyde Measurements

The Malondialdehyde (MDA) was measured according to a previous study [29]. 4-week-old rosette leaves of 10 plants treated by dark submergence were weighed and pulverized in 5% trichloroacetic acid buffer, and then mix the supernatant with 6.7% thiobarbituric acid and 5% trichloroacetic acid buffer. The materials were further incubated at 100°C for 0.5 hr, and then cooled to the room temperature. The absorbance was measured at 532, 450, and 600 nm with a spectrophotometer plate reader.

ChIP and ChIP-sequencing

Samples of 14-day-old seedlings of *WRKY33OE1* and *WRKY33OE2* plants were fixed using 1% formaldehyde and prepared for chromatin immunoprecipitation assays, as previously described [30]. The DNA-protein complexes were extracted from rosette leaves (2 g pooled leaf materials) of 4-week-old *35S:FLAG-WRKY33 OE1* and *OE2* transgenic plants, and pulled down using anti:FLAG antibody (Sigma-Aldrich F1084) and protein A Agarose beads following the ChIP protocol [31]. The immunoprecipitated DNA fragments were dissolved in 40 μl ddH2O and then sent to the BGI (Shenzhen, China) company for the following experiment. 10% of the total DNA-protein complexes before the immunoprecipitation were used as the input DNA.

ChIP-Seq service was performed by BGI company (Shenzhen, China). The DNA is combined with End Repair Mix and incubated at 20°C for 30 min. We further purified the end-repaired DNA with QIAquick PCR Purification Kit (Qiagen), and added A-Tailing Mix and incubated at 37°C for 30 min. We combined the purified Adenylate 3 'Ends DNA, Adapter and Ligation Mix and incubated the ligation reaction at 20°C for 15 min. We purified the Adapter-ligated DNA with the QIAquick PCR Purification Kit. We conducted several rounds of PCR amplification with PCR Primer Cocktail and PCR Master Mix to enrich the Adapter-ligated DNA fragments. Then the PCR products are selected (about 100–300 bp, including adaptor sequence) by running a 2% agarose gel to recover the target fragments. We purified the gel with QIAquick Gel Extraction kit (QIAGEN). The final library was quantitated in two ways: determining the average molecule length and sample integrity and purity using the Agilent 2100 bioanalyzer instrument (Agilent DNA 1000 Reagents) and quantifying the library by real-time quantitative PCR (QPCR). The double stranded PCR products were heat-denatured and circularized by the splint oligo sequence. The single strand circle DNA (ssCir DNA) was formatted as the final library. Library was qualified by Qubit ssDNA kit. The sequencing was performed with the BGISEQ-500 sequencing system, featured by combinatorial probe-anchor synthesis (cPAS) and DNA Nanoballs (DNB) technology for superior data quality (BGI-Shenzhen, China).

The raw sequencing image data were examined by the Illumina analysis pipeline. ChIP-Seq reads were aligned to the *Arabidopsis* reference genome (TAIR10) by Bowtie [32] with at most 2 mismatches. The input group was used as a control. The results were visualized with IGV software. Reads that appeared

more than twice at the same position on the same strand were discarded to remove PCR duplication. MACS2 (Model-based Analysis of ChIP-Seq) [33] was used to identify peaks using a q-value cutoff of 0.05.

Motif analysis

To identify possible binding motif of the WRKY33 transcription factor, the ChIP peak sequences were subjected to MEME (Multiple EM for Motif Elicitation)-ChIP [19]. The MEME-ChIP program uses two ab initio motif discovery algorithms: MEME [20], and DREME (Discriminative Regular Expression Motif Elicitation) [34], which uses regular expressions to search for short eukaryotic TF motifs that are missed by MEME.

Gene function of WRKY33 TF target genes

In order to determine the putative functions of the target gene *WRKY33*, all identifed genes with ChIP-Seq peaks in the upstream promoter region or the potential regulatory region downstream were subjected to annotation of the categories of ontological genes (GO) [21]. The default Fisher's Exact Test and Benjamini-Yekutieli multiple test correction methods [35] were used to generate *p*-values for statistical significance and corresponding False Discovery Rate (FDR) values.

RNA extraction and quantification

Total RNA was isolated using the Biospin Plant Total RNA Extraction kit according to the user manual (Bioer Technology; Hangzhou, China), from the pooled three-week old rosette leaves of Col and *35S:FLAG-WRKY33* plants, and 1–2 µg total RNA was used for reverse transcription, using the PrimeScript RT reagent kit (Takara Cat# RR047A). A QuantiNova SYBR Green PCR Kit was used for qPCR reactions with qPCR-specific primers. The expression levels of putative target genes were compared with *ArabidopsisACTIN* genes.

Abbreviations

At: *Arabidopsis thaliana*; ChIP: chromatin immunoprecipitation; DREME: Discriminative Regular Expression Motif Elicitation; GO: Gene Ontology; MEME: Multiple EM for Motif Elicitation; RT: reverse transcriptase; Seq: sequencing; TF: transcription factor.

Declarations

Accession numbers

Sequence data generated in this article can be found in the GenBank/EMBL libraries under the following accession numbers: *WRKY33* (AT2G38470), AT2G35736, AT1G66810, AT2G47090, AT3G12120, and *ACTIN2* (AT3G18780).

Acknowledgements

This research was equally supported by the National Natural Science Foundation of China (31870244).

Authors' contributions

LHH and XZX designed the experiments; ZJL, LB, SY, CY and FJ performed the experiments and analyzed the data; LHH and XZX wrote the article, LJQ and MT revised the article.

Funding

The Fundamental Research Funds for the Central Universities (grant No. SCU2019D013).

Availability of data and materials

All data generated are included in this published article and its supplementary files.

Ethics approval and consent to participate

The original seeds of *Arabidopsis thaliana* were kept in our lab in the Key Laboratory for Bio-resources and Eco-environment, College of Life Science, Sichuan University. The experimental methods conducted in this study complied with current Chinese laws and regulations. The trade name, company name, or company name used in this publication is to provide readers with information and convenience. Such use does not constitute an official endorsement or endorsement of any product or service by the Ministry of Agriculture or Agricultural Research Service Department of China, does not exclude other suitable products or services.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Armstrong WDJAiBR. Aeration in Higher Plants. Advances in Botanical Research. 1979; 7:225-332.
2. Visser EJW, Voesenek LACJ, Vartapetian BB, Jackson MB. Flooding and Plant Growth: Academic Press. 1984.
3. Kennedy RA, Rumpho ME, Physiology TCFJP. Anaerobic Metabolism in Plants. Plant Physiology. 1992; 100(1):1-6.
4. Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ESJPC. Expression Profile Analysis of the Low-Oxygen Response in Arabidopsis Root

Cultures. Plant Cell. 2002; 14(10):2481-2494.

5. Licausi F, Kosmacz M, Weits DA, Giuntoli B, Giorgi FM, Voesenek LACJ, Perata P, van Dongen JT. Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. Nature. 2011; 479(7373):419-422.
6. Dissmeyer NJARoPB. Conditional Protein Function via N-Degron Pathway–Mediated Proteostasis in Stress Physiology. Annual Review of Plant Biology. 2019; 70(1):83-117.
7. Cheatle Jarvela AM, Hinman VFJE. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. Evodevo. 2015; 6(1):1-11.
8. Moreno-Risueno MÁ, Martínez M, Vicente-Carbajosa J, Carbonero PJMG, Genomics. The family of DOF transcription factors: from green unicellular algae to vascular plants. Molecular Genetics. 2007; 277(4):379-390.
9. Martinez-Castilla LP, Alvarez-Buylla ERJPotNAoeotUSoA. Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny. Proceedings of the National Academy of ences of the United States of America. 2003; 100(23):13407-13412.
10. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 2000; 5(5):199-206.
11. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. Trends Plant Sci. 2010; 15(5):247-258.
12. Bakshi M, Oelmüller RJPSB. WRKY transcription factors: Jack of many trades in plants. Plant Signal Behav. 2014; 9(2):e27700.
13. Krishnamurthy P, Vishal B, Wan JH, Lok FCJ, Kumar PP. Regulation of CYP94B1 by WRKY33 controls apoplastic barrier formation in the roots leading to salt tolerance. Plant Physiol. 2020.
14. Rajappa S, Krishnamurthy P, Kumar PPJFiPE. Regulation of AtKUP2 Expression by bHLH and WRKY Transcription Factors Helps to Confer Increased Salt Tolerance to Arabidopsis thaliana Plants. Frontiers in Plant Ence. 2020; 11:1311.
15. Sun Y, Liu Z, Guo J, Zhu Z, Sun XJB, Communications BR. WRKY33-PIF4 loop is required for the regulation of H2O2 homeostasis. Biochem Biophys Res Commun. 2020; 527(4):922-928.
16. Barco B, Kim Y, Clay N. Expansion of a core regulon by transposable elements promotes Arabidopsis chemical diversity and pathogen defense. Nat Commun. 2019; 10(1):3444.
17. Tang H, Bi H, Liu B, Lou S, Yan S, Tong S, Chen N, Jiang Y, Liu J, Liu HJNP. WRKY33 interacts with WRKY12 protein to up-regulate RAP2.2 during submergence induced hypoxia response in Arabidopsis thaliana. New Phytologist. 2020; (1469-8137).
18. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based Analysis of ChIP-Seq (MACS). Methods in Molecular Biology. 2014; 1150:81-95.
19. Machanick, Philip, Bailey, Timothy, Bioinformatics LJ. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011; 27(12):1696-1697.

20. Bailey TLJB. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011; 27(12):1696-1697.
21. BeiBbarth T, Speed TPJB. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics. 2004; 20(9):1464-1465.
22. Perata P, Voesenek LACJJTiPS. Submergence tolerance in rice requires Sub1A, an ethylene-response-factor-like gene. Trends in Plant Science. 2007; 12(2):43-46.
23. Datta R, Kumar D, Sultana A, Hazra S, Bhattacharyya D, Chattopadhyay SJPP. Glutathione Regulates 1-Aminocyclopropane-1-Carboxylate Synthase Transcription via WRKY33 and 1-Aminocyclopropane-1-Carboxylate Oxidase by Modulating Messenger RNA Stability to Induce Ethylene Synthesis during Stress. Plant Physiology. 2015; 169(4):2963.
24. Jiang Y, Deyholos MKJPMB. Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Molecular Biology. 2009; 69(1):91-105.
25. Wang J, Zhou L, Shi H, Chern M, Yu H, Yi H, He M, Yin J, Zhu X, Li Y, Li W, Liu J, Wang J, Chen X, Qing H, Wang Y, Liu G, Wang W, Li P, Wu X, Zhu L, Zhou JM, Ronald PC, Li S, Li J, Chen X. A single transcription factor promotes both yield and immunity in rice. Science. 2018; 361(6406):1026-1028.
26. Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou G, Whitham SAJPC. Expression Profile Matrix of Arabidopsis Transcription Factor Genes Suggests Their Putative Functions in Response to Environmental Stresses. Plant Cell. 2002; 14(3):559-574.
27. Trevaskis B, Watts RA, Andersson CR, Llewellyn DJ, Hargrove MS, Olson JS, Dennis ES, Peacock WJJPotNAoSotUSoA. Two hemoglobin genes in Arabidopsis thaliana: The evolutionary origins ofleghemoglobins. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94(22):12230-12234.
28. Xiuren Z, Rossana H, Shih-Shun L, Qi-Wen N, Nam-Hai CJNP. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nature Protocol. 2006; 2(1):1-6.
29. Yuan L-B, Dai Y-S, Xie L-J, Yu L-J, Zhou Y, Lai Y-X, Yang Y-C, Xu L, Chen Q-F, Xiao SJPP. Jasmonate Regulates Plant Responses to Reoxygenation through Activation of Antioxidant Synthesis. Plant Physiology. 2017; 173(3):1864.
30. Martin RC, Vining K, Dombrowski JEJBPB. Genome-wide (ChIP-seq) identification of target genes regulated by BdbZIP10 during paraquat-induced oxidative stress. Bmc Plant Biology. 2018; 18(1):58.
31. Bowler C, Benvenuto G, Laflamme P, Molino D, Probst AV, Tariq M, Paszkowski JJPJ. Chromatin techniques for plant cells. Plant Journal. 2010; 39(5):776-789.
32. Langmead B, Trapnell C, Pop M, Salzberg SLJGB. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology. 2009; 10(3):R25.
33. Feng J, Liu T, Zhang YJCPB. Using MACS to identify peaks from ChIP-Seq data. Curr Protoc Bioinformatics. *2011; Chapter 2(1):Unit 2.14.*
34. Bailey TLJB. DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics. 2011; 27(12):1653-1659.

35. Benjamini Y, Yekutieli DJAoS. The control of the false discovery rate in multiple testing under dependency. Annals of Stats. 2001; 29(4):1165–1188.

Figures

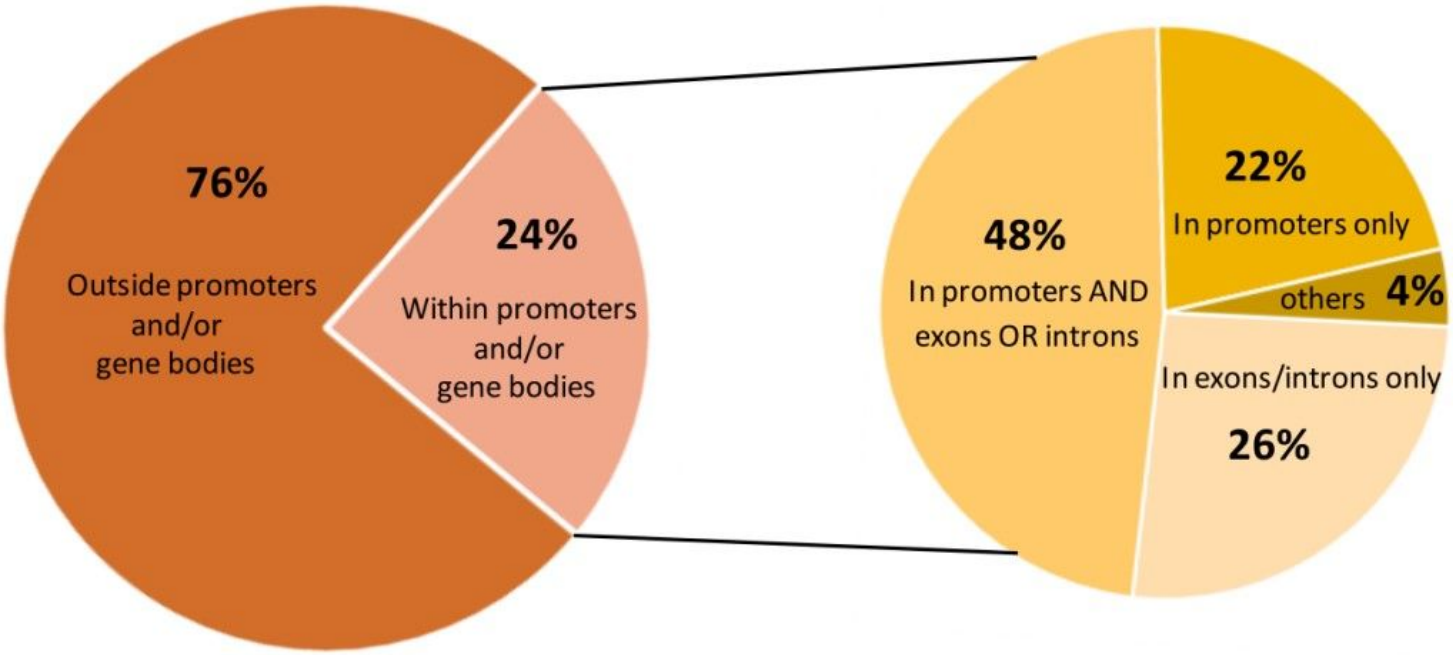

Figure 1

Distribution of ChIP peaks in the genome. Percentage of peaks that reside 2 kb upstream of the transcription start site or 2 kb downstream of the stop codon (gene body), and location of the peaks in the gene bodies.

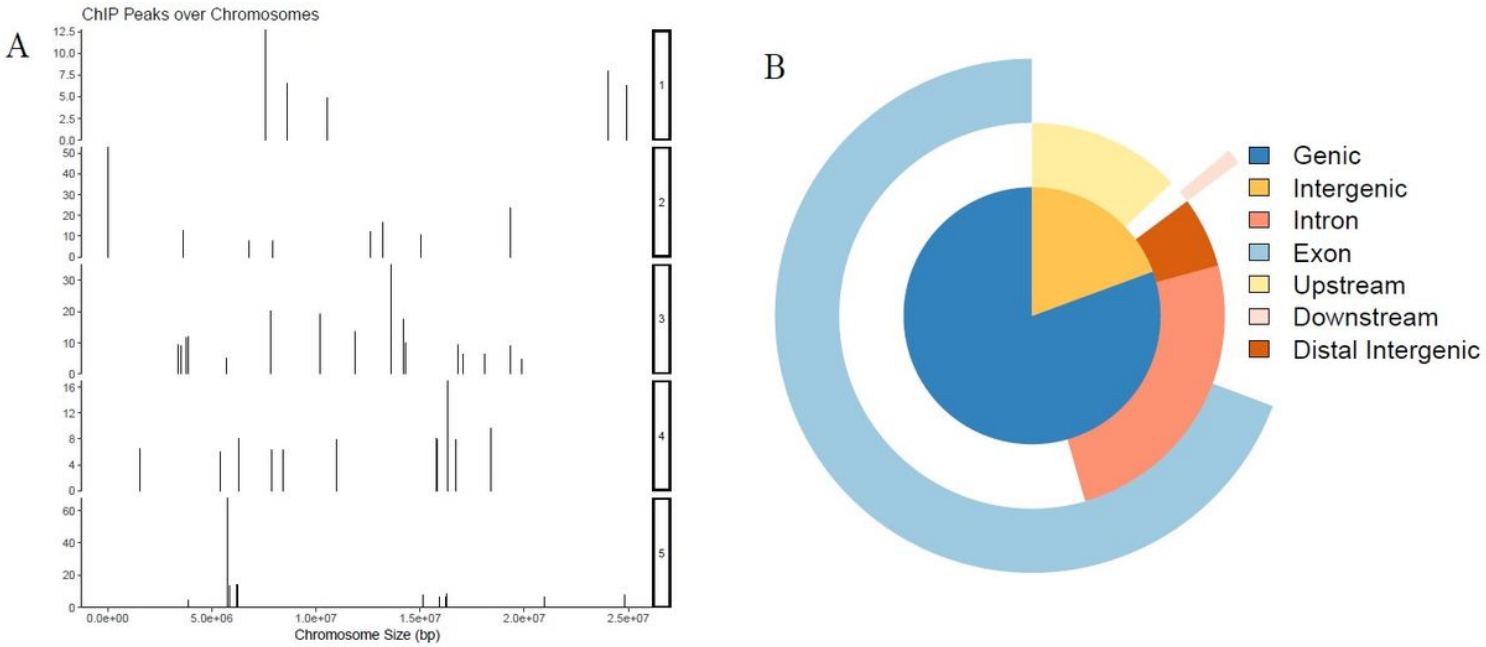

Figure 2

The location of all ChIP peaks over chromosome. (A) ChIP peaks coverage plot: the right ordinate represents the chromosome, the left ordinate represents the size of the peak, and the abscissa represents the size of the chromosome. (B) Genomic Annotation by vennpie. Visually shows the full annotation with their overlap.

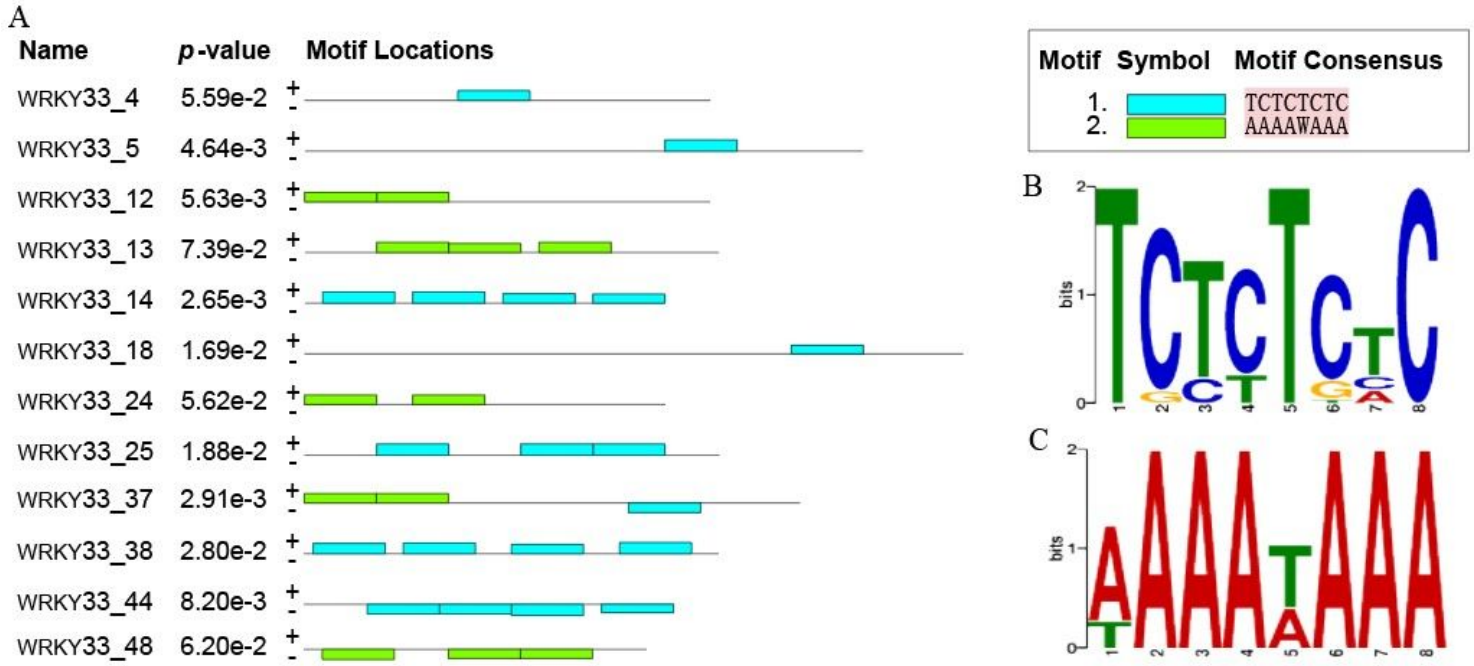

Figure 3

Genome-wide distribution of WRKY33 binding sites in the Arabidopsis genome identified by ChIP-Seq. (A) MEME-CHIP analysis of WRKY33 motif. Arabidopsis reference genome (TAIR10) by Bowtie, Among the two motifs identified by MEME, ChIP peaks and p value and locations where two motifs are located. (B-C) The two most representative motif patterns.

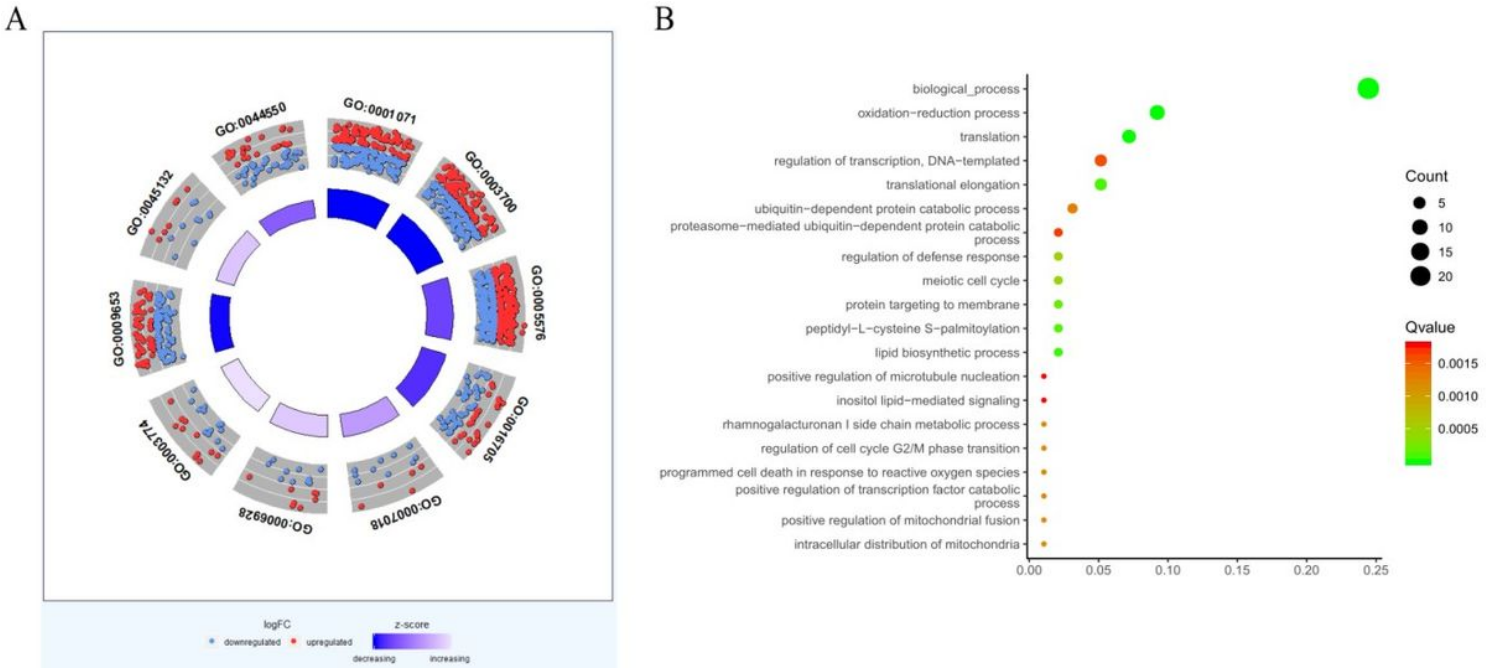

Figure 4

(A) Top 10 significantly enriched GO biological processes of WRKY33BGHs. Red and blue dots indicate up-regulated DEGs and down-regulated DEGs enriched in the term respectively, and a z-score indicated in the inner quadrangle. (B) The results of the top 20 extremely significant enrichments indicate that the gene ontology categories for biological processes includes the oxidation-reduction process and programmed cell death in response to reactive oxygen species.

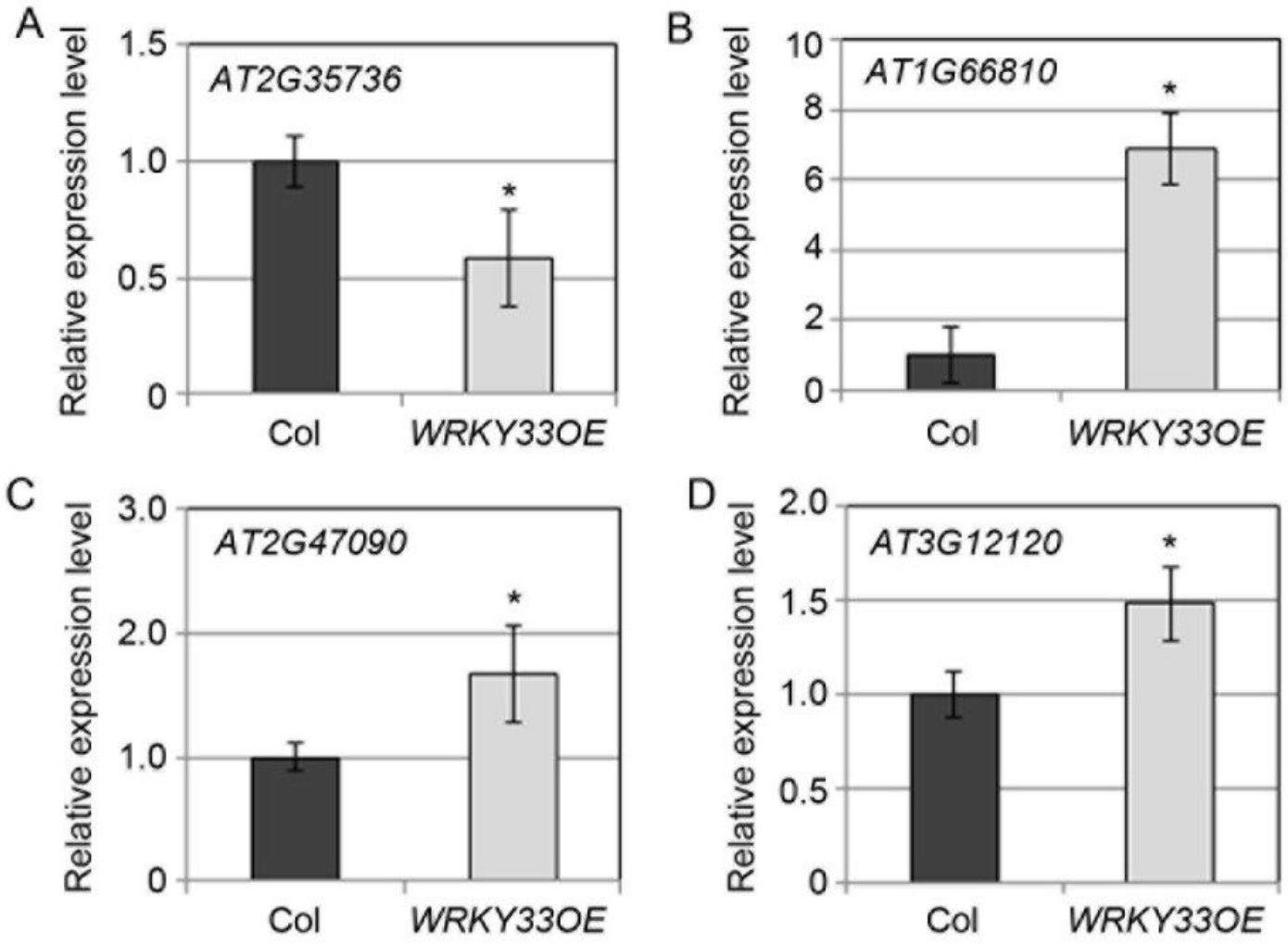

Figure 5

Expression analysis of genes containing the “TC box” in Col and WRKY33OE plants after submergence treatment. (A) AT2G35736 gene is downregulated by WRKY33 upon submergence treatment. (B-D) AT1G66810, AT2G47090 and AT3G12120 genes are upregulated by WRKY33 upon submergence treatment. Three independent biological replicates were used. Data are average values ±SD (n=3) of 3 biological replicates. *($p < 0.05$) indicates significant difference from Col.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFig.1.docx
- SupplementalFig.2.docx
- Additionalfile1.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx