Plant Physiology and Biochemistry 103 (2016) 71–83

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Genome-wide identification and characterization of WRKY transcriptional factor family in apple and analysis of their responses to waterlogging and drought stress

Dong Meng [a, 1], Yuanyuan Li [a, b, 1], Yang Bai [c], Mingjun Li [a, d], Lailiang Cheng [a, *]

[a] Department of Horticulture, Cornell University, 134A Plant Science, Ithaca, NY 14853, USA
[b] College of Horticulture Science and Engineering, Shandong Agricultural University, Taian 271018, Shandong, China
[c] Boyce Thompson Institute for Plant Research, Ithaca, NY 14853, USA
[d] College of Horticulture, Northwest A&F University, Yangling, Shaanxi, 712100, China

ARTICLE INFO

Article history:
Received 31 December 2015
Received in revised form
5 February 2016
Accepted 5 February 2016
Available online 2 March 2016

Keywords:
WRKY
Apple
Transcriptional factor
Water stress
Drought stress

ABSTRACT

As one of the largest transcriptional factor families in plants, *WRKY* genes play significant roles in various biotic and abiotic stress responses. Although the *WRKY* gene family has been characterized in a few plant species, the details remain largely unknown in the apple (*Malus domestica* Borkh.). In this study, we identified a total of 127 *MdWRKYs* from the apple genome, which were divided into four subgroups according to the WRKY domains and zinc finger motif. Most of them were mapped onto the apple's 17 chromosomes and were expressed in more than one tissue, including shoot tips, mature leaves, fruit and apple calli. We then contrasted *WRKY* expression patterns between calli grown in solid medium (control) and liquid medium (representing waterlogging stress) and found that 34 *WRKY* genes were differentially expressed between the two growing conditions. Finally, we determined the expression patterns of 10 selected *WRKY* genes in an apple rootstock, G41, in response to waterlogging and drought stress, which identified candidate genes involved in responses to water stress for functional analysis. Our data provide interesting candidate *MdWRKYs* for future functional analysis and demonstrate that apple callus is a useful system for characterizing gene expression and function in apple.

1. Introduction

Since the SPF1 was first identified in sweet potato as a DNA binding protein (Ishiguro and Nakamura, 1994), abundant *WRKY* transcriptional factors have been found throughout the green lineage (green algae and land plants) (Ulker and Somssich, 2004). WRKY proteins constitute a large transcriptional factor family in plants (Eulgem et al., 2000; Rushton et al., 2010). Its members are characterized by the WRKY motif, which includes a highly conserved WRKYGQK signature peptide sequence together with a zinc finger motif (Rushton et al., 1995; Eulgem et al., 2000). Based on the number of WRKY domains and the type of zinc finger motifs, they can be divided into four groups (Wang et al., 2014). The proteins with two WRKY domains compose group I, whereas group II and group III members have only a single WRKY domain, most followed by a novel zinc-finger motif C2H2 (C-X_{4-5}-C-X_{22-23}-H-X-H) or C2HC (C-X_7-C-X_{23}-H-X-C), respectively (Eulgem et al., 2000). Group IV proteins contain the WRKY domain but lack a complete zinc-finger structure in the C-terminus. Moreover, most of the WRKY proteins bind to the cis-acting W-box elements ((T)TGAC(C/T)) in the target promoter region to regulate gene expression and participate in the cell signaling network through the conserved WRKY domain (Eulgem et al., 2000; Wang et al., 2014).

As one of the largest transcriptional factor families in higher plants, WRKY proteins have been found to play essential roles in various biotic and abiotic stresses. They are involved in the plant immune system mediated by hormones (jasmonic acid and salicylic acid) that can respond to attacks by pathogens, such as bacteria, viruses and fungi (Mukhtar et al., 2008; Yang et al., 1999; Marchive et al., 2007; Eulgem and Somssich, 2007). In Arabidopsis, AtWRKY70 is required for R gene-mediated pathogen resistance, determining the balance between the SA- and JA-dependent defense system (Li et al., 2006; Knoth et al., 2007). Many *WRKY* genes

* Corresponding author.
E-mail address: LC89@Cornell.edu (L. Cheng).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.plaphy.2016.02.006
0981-9428/

also act as negative regulators of defense signaling; for example, *AtWRKY38* and *AtWRKY62* contribute negatively to basal resistance toward bacterial pathogens (Kim et al., 2008). Another important role of *WRKY* genes is their participation in many complex signaling progresses during plant abiotic stress responses, such as drought, salinity, heat and freezing stresses (Hara et al., 2000; Pnueli et al., 2002; Huang and Duman, 2002). However, relative to the large number of WRKY proteins in different plants and their diverse roles under complex environmental stimulations, there is much less information available in relation to abiotic stress research. Understanding WRKY protein functions in abiotic stress remained a great challenge until recently, when large-scale sequencing, genome-wide analysis and transcription profiling became available for characterizing genes in various plants. The *WRKY* genes in many plants have been identified, and their functions under abiotic stresses were inferred via detection of their transcript levels.

Although WRKY proteins have been explored in a few plant species, including *Arabidopsis*, rice (*Oryza sativa*), maize, cucumber and cotton, and grape (Rushton et al., 2010; Tripathi et al., 2012; Wei et al., 2012), the WRKY family in the apple remains uncharacterized. As a woody perennial species, apple is among the most valuable fruit crops in the world and often suffers from various environmental stresses. Based on this consideration, we performed a genome-wide identification of apple *WRKY* genes using the apple genome sequence information (GDR, www.rosaceae.org). A phylogenetic tree was constructed for the identified WRKY proteins. Then, the *WRKY* gene expression patterns in different apple tissues were detected through our RNA-seq database. Additionally, we contrasted the expression patterns of apple calluses grown under two different conditions (solid and liquid medium) as determined by RNA-seq. Finally, we used real time PCR to detect the expression levels of 10 selected *WRKY* genes in apple rootstock G41 trees under waterlogging and drought stresses.

2. Materials and methods

2.1. Database search and sequence annotation of the WRKY family in apple

The sequences of Arabidopsis WRKY proteins were downloaded from the Arabidopsis genome (TAIR, http://www.Arabidopsis.org/). Two different approaches were used to search for *Malus domestica* (Md) WRKY transcription factors. First, we selected the WRKY conserved domains (WRKYGQK signature protein sequence) of Arabidopsis WRKY proteins as a query sequence to search the apple genome sequence database on the website GDR (Genome Database for *Rosaceae*, https://www.rosaceae.org/). A BLASTp search was performed, and sequences were selected for further analysis where the E value was less than −10. After the batch BLAST searches, we checked the WRKY transcription factors in the PlantTFDB database (http://planttfdb.cbi.pku.edu.cn/) to retrieve any additional *WRKY* genes. We also used the coding sequences (CDS) to perform blast searches against the Phytozome database (http://phytozome.jgi.doe.gov/pz/portal.html#). Any additional *WRKY* genes were retrieved for further analysis. To further verify the reliability of those WRKY candidate sequences, the SMART database (http://smart.embl-heidelberg.de/) and NCBI Conserved Domains (CD)-search tools (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) were used to confirm each predicted MdWRKY protein as a member of WRKY family. Redundant sequences or sequences lacking the WRKY domain were removed.

2.2. Phylogenetic tree construction and sequence analysis

AtWRKY proteins were downloaded from the Arabidopsis genome (TAIR, http://www.Arabidopsis.org/). Phylogenetic analysis was conducted using MEGA version 6.0 (http://www.megasoftware.net/) with the neighbor-joining method, and bootstrap values were calculated using 1000 iterations. Eight Arabidopsis WRKY domains from different WRKY subgroups were used as references to categorize the WRKY proteins from apple using the MUSCLE algorithm integrated in the MEGA6.0 package (Tamura et al., 2013). The GenBank accession numbers of those AtWRKYs are AtWRKY1: ABJ17102, AtWRKY9: At5g08670, AtWRKY14: AAP21276.1, AtWRKY18: AAM78067, AtWRKY21: AAB63078.1, AtWRKY31: AEE84546.1, AtWRKY41: AEE82969, AtWRKY45: ABD57509.1. The exon and intron organizations of *MdWRKY* genes were determined by comparing predicted CDS with their corresponding genomic sequences using GSDS software (http://gsds.cbi.pku.edu.cn/). Structural motif annotation was performed using the MEME program (http://meme-suite.org/). The physical distribution of the *MdWRKY* genes on chromosomes was drawn using MapChart based on gene position in the apple genome. The chromosome locations of the genes used in this paper were based on markers supplied by GDR website.

2.3. Plant materials and growth conditions

Calli of the apple variety 'Orin' were maintained at 3-week intervals on Murashige and Skoog medium containing 1.5 mM 2,4-D and .4 mM 6-benzylaminopurine in the dark at 25 °C, unless stated otherwise. The solid medium contained 8 g/L agar. For the calli grown in liquid medium, agar was omitted and the calli were cultured in flasks on a shaker at 120 rpm/min. Callus samples were taken when they were 15 days old and frozen in liquid nitrogen for later use.

Plants of 'G.41', an apple (*M. domestica* Borkh.) rootstock, propagated via tissue culture, were used in the water stress experiment. They were grown in 10 cm (diameter) × 9 cm (height) pots containing Cornell mix and sand in a 3:1 volumetric ratio in a greenhouse with a day/night temperature of 25 °C/15 °C and a minimum photosynthetically active radiation of 250 μmol m^{-2} s^{-1} supplemented with high pressure sodium lights in a photoperiod of 14 h at Cornell University in Ithaca, New York, USA. When they reached about 20 cm tall, uniform plants were selected for water stress treatments. For the well-watered control, pots were watered to drip each day. For the waterlogging treatment, pots were submerged in water for 7 days. For the drought stress treatment and recovery, water was withheld for 5 days and then the plants were watered for 2 days the same way as the well-watered control. Each treatment was replicated 5 times with 3 plants per replicate in a completely randomized design. Root and leaf samples were taken at 5 h into the photoperiod on Day 0, 1, 3, 5, and 7 after the initiation of the treatments and frozen in liquid nitrogen for later use.

2.4. Library construction and RNA-Seq analysis

Total RNA was extracted using the modified CTAB method (Gasic et al., 2004) and treated with DNase (Bio-Rad Laboratories, Hercules, CA, USA) prior to mRNA isolation. Libraries corresponding to five biological replicates from solid and liquid cultured apple calluses were constructed following a strand-specific RNA-seq library preparation protocol as previously described (Zhong et al., 2011), and multiplexed into 1 lane for sequencing at the Genomics Facility, Institute of Biotechnology, Cornell University (Ithaca, NY, USA). Sequencing was performed using an Illumina HiSeq 2000 sequencer with 50 bp single end sequencing (Illumina, San Diego, CA).

The RNA-seq reads were processed by removing barcode and adaptor sequences first, followed by alignments to rRNA database