Postprint: Bioinformatics Analysis of the WRKY Gene Family in *Coffea canephora*

Authors: Fu Chun, Tang Xue, Yang Yaojun, Jiang Na

Date: 2020-08-02T00:00:00+00:00

Abstract

WRKY transcription factors are indispensable components of plant signaling networks and, as one of the largest transcription factor families in plants, play important roles in various plant stress responses. Using bioinformatics methods, a detailed analysis of the physicochemical properties and molecular evolution of the WRKY protein family in Coffea canephora was conducted. The results showed that the number of amino acids in CcWRKY proteins ranged from 103 to 994, none contained signal peptides, and they were predicted to be non-secretory proteins. Their secondary structures were dominated by random coils as the primary structural element, while tertiary structures were mainly classified into 6 categories, with the D class structure—represented by CcWRKY61, CcWRKY72, CcWRKY6, and CcWRKY7—being the most stable. Conserved domain and phylogenetic tree analyses indicated that the WRKY gene family in Coffea canephora comprised 49 members, of which 10 members were assigned to WRKY Group , 34 members to WRKY Group , and 5 members to WRKY Group . Phylogenetic analysis of the WRKY1 gene from Coffea canephora with other species revealed that its WRKY1 showed the closest relationship with tobacco and the most distant relationship with African oil palm (Elaeis guineensis), suggesting that the WRKY1 protein is relatively conserved during biological evolution. The findings of this study can provide a valuable reference for in-depth investigation of the molecular functions of the WRKY gene family in Coffea canephora and hold significant practical importance for further exploring the functions, evolution, and molecular breeding of WRKY genes in Coffea canephora.

Full Text

Preamble

Bioinformatics Analysis of the WRKY Gene Family in *Coffea canephora*

FU Chun1,2*, TANG Xue1,2, YANG Yaojun1,2, JIANG Na3

1 Bamboo Disease and Pest Control and Resources Development Key Laboratory of Sichuan Province, Leshan Normal University, Leshan 614000, Sichuan, China
2 College of Life Science, Leshan Normal University, Leshan 614000, Sichuan, China
3 College of Mathematics and Information Science, Leshan Normal University, Leshan 614000, Sichuan, China

Abstract

WRKY transcription factors represent an indispensable component of plant signaling networks. As one of the largest families of transcription factors in plants, they play crucial roles in various stress responses. This study employed bioinformatics methods to conduct a comprehensive analysis of the physicochemical properties and molecular evolution of the WRKY protein family in *Coffea canephora*. The results revealed that the amino acid length of CcWRKY proteins ranged from 103 to 994 residues, with no signal peptides detected, suggesting they are non-secretory proteins. Random coil constituted the predominant secondary structural element, while tertiary structures were primarily divided into six classes, with the D-class structure—represented mainly by CcWRKY61, CcWRKY72, CcWRKY6, and CcWRKY7—exhibiting the highest stability. Conserved domain and phylogenetic tree analyses indicated that the *C. canephora* WRKY gene family comprises 49 members, including 10 classified as WRKY Group I, 34 as Group II, and 5 as Group III. Phylogenetic analysis of the *C. canephora* WRKY1 gene with other species demonstrated its closest relationship with tobacco and most distant relationship with oil palm (*Elaeis guineensis*), suggesting that the WRKY1 protein is relatively conserved during biological evolution. These findings provide valuable insights for in-depth investigation of the molecular functions of the *C. canephora* WRKY gene family and hold significant practical importance for further exploration of WRKY gene function, evolution, and molecular breeding in coffee.

Keywords: *Coffea canephora*, WRKY, transcription factor, phylogenetic evolution, bioinformatics

Introduction

Coffea canephora (robusta coffee) ranks first among the three major beverage crops in terms of average annual production, output value, and consumption (Yan et al., 2018). As an important stimulant beverage with a history spanning hundreds of years, it represents a crucial commodity in international trade (Davis et al., 2011; Monente et al., 2015). *C. canephora* is a small tree or

shrub belonging to the Rubiaceae family. The genus *Coffea* comprises over 100 species, with commercial cultivation primarily focusing on two species: *C. arabica* (arabica coffee) and *C. canephora*, accounting for approximately 60% and 40% of global coffee production, respectively. In 2014, Hainan Province primarily cultivated *C. canephora*, with coffee plantations covering 0.47% of the national total (Rodrigues et al., 2015; Yan et al., 2018). Robusta coffee thrives in high-temperature, high-rainfall environments; however, its cultivation frequently encounters stresses from diseases, drought, and low temperature, which severely affect growth and development, causing substantial production losses (Yang, 2014).

In plants, WRKY proteins constitute a ubiquitous superfamily of transcription factors that ranks among the most extensively studied regulatory families in plant science, belonging to the ten major superfamilies of transcription factors in higher plants (Zheng et al., 2018). The WRKY nomenclature derives from the conserved seven-amino-acid sequence (WRKYGQK) present at the N-terminus of all family members. Based on this characteristic, the plant WRKY gene superfamily is classified into three groups: I, II, and III (Zheng et al., 2018). Group I contains two WRKY domains and a C2H2 zinc finger motif (C-X4-5-C-X22-23-H-X1-H, where X represents any amino acid). Group II features one WRKY domain and a C2H2 zinc finger structure (C-X4-5-C-X22-23-H-X1-H). Group III comprises one WRKY domain and a C2HC-type zinc finger structure (C-X7-C-X23-H-X1-C) (Bakshi et al., 2014).

Ishiguro & Nakamura (1994) first cloned the initial WRKY gene from sweet potato (*Ipomoea batatas*) and identified a DNA-binding protein designated SPF1 transcription factor. Subsequently, numerous WRKY transcription factors have been cloned and identified across various crops, including castor bean (56 WRKY genes) (Zou, 2013), kiwifruit (89 WRKY genes) (Bao et al., 2018), tomato (81 WRKY genes) (Zhang, 2017), grape (56 WRKY genes) (Su et al., 2019), and jujube (92 WRKY genes) (Xue et al., 2018).

WRKY transcription factors represent a plant-specific family involved in regulating seed dormancy release, metabolism, and hormone signal transduction (Xie et al., 2016). They function as transcriptional activators or repressors in numerous plant developmental and physiological processes (Rushton et al., 2010; Rushton et al., 2012; Ding et al., 2015). Research has demonstrated that WRKY gene families play critical roles in responding to abiotic stresses such as salt, high temperature, low temperature, and drought (Jiang et al., 2006; Wu et al., 2009; Li et al., 2019). Additionally, WRKY transcription factors participate in plant growth and developmental processes, including fruit ripening (Xu et al., 2004), embryogenesis (Lagacé et al., 2004), and senescence (Robatzek et al., 2010; Zhang et al., 2019). In kiwifruit, 33 WRKY genes exhibit significant expression across four organs (roots, leaves, flowers, and fruits) (Bao et al., 2018). In apple, 12 MdWRKY genes are expressed in roots, stems, leaves, flowers, and fruits, potentially regulating apple growth and development (Zhang et al., 2019). Studies on WRKY transcription factors in tobacco 'Yunyan

87' leaf senescence revealed that NtWRKY11, NtWRKY31, NtWRKY40, and NtWRKY51 are associated with both stress responses and leaf senescence (Gu et al., 2015). To date, while the TPS gene family has been systematically analyzed in *C. canephora* transcription factors (Cheng et al., 2016), WRKY gene family research remains limited. Dong et al. (2019) identified 49 CcWRKYs in the *C. canephora* genome and investigated their responses to cold stress. This study employs bioinformatics approaches to predict and analyze the amino acid composition, transmembrane domains, conserved motifs, protein tertiary structures, and phylogenetic relationships with tobacco, kiwifruit, rice, and other species, providing a theoretical foundation for further functional exploration of WRKY transcription factors related to biotic stress.

1. Materials and Methods

1.1 Experimental Materials

The nucleotide and amino acid sequences of WRKY transcription factor families from 12 species—*Coffea canephora*, rice (*Oryza sativa* subsp. *indica*), maize (*Zea mays*), kiwifruit (*Actinidia chinensis*), tobacco (*Nicotiana tabacum*), Chinese chestnut (*Castanea mollissima*), oil palm (*Elaeis guineensis*), banana (*Musa acuminata*), moso bamboo (*Phyllostachys heterocycla*), jujube (*Ziziphus jujuba*), apple (*Malus domestica*), and *Arabidopsis thaliana*—were obtained from the Plant Transcription Factor Database (http://planttfdb.cbi.pku.edu.cn/index.php, version V5.0). The correspondence between *C. canephora* WRKY (CcWRKY) gene family members and *Arabidopsis* WRKY protein sequences was identified using the Blastp program (Table 1).

1.2.1 Analysis of Physicochemical Properties of WRKY Family-Encoded Proteins

The physicochemical properties of CcWRKY proteins, including amino acid length, grand average of hydropathicity, and numbers of positively and negatively charged residues, were analyzed using the ProtParam online tool. Signal peptide prediction and subcellular localization analysis of 49 *C. canephora* WRKY gene family amino acid sequences were performed using SignalP-5.0 and CELLO v2.5 software (Yu et al., 2006; Sha et al., 2019), respectively.

1.2.2 Prediction of Transmembrane Structure, Hydrophobicity/Hydrophilicity, Phosphorylation Sites, and Conserved Motifs

The transmembrane structure of *C. canephora* WRKY proteins was analyzed using the TMHMM Server v.2.0 online tool (Zhou et al., 2018). Hydrophobicity/hydrophilicity prediction was completed via the ProtScale online server. Potential phosphorylation sites in *C. canephora* WRKY protein sequences were predicted using NetPhos3.1 Server (http://www.cbs.dtu.dk/services/NetPhos/). Conserved motifs in *C. canephora* WRKY gene family members were searched

using MEME 5.0.4 online tool with parameters set to identify 10 conserved motifs and default settings for other parameters.

1.2.3 Prediction of Protein Secondary and Tertiary Structures

Secondary structures of 49 *C. canephora* WRKY proteins were predicted and analyzed using the SOPMA online server (Zhang et al., 2018; Zhou et al., 2018). Tertiary structures of 49 WRKY proteins were predicted using the Phyre 2 web portal (Kelley et al., 2015).

1.2.4 Construction of Phylogenetic Trees for *C. canephora* WRKY Gene Family

MEGA 7.0 software (Wang et al., 2019) was used for multiple sequence alignment of CcWRKY protein sequences and construction of phylogenetic trees. Phylogenetic trees were also constructed for WRKY1 protein sequences from *C. canephora* and 11 other species including *Arabidopsis*, tobacco, kiwifruit, Chinese chestnut, apple, and maize. The neighbor-joining method was employed (Saitou et al., 1987; Kumar et al., 2018), with bootstrap resampling performed 1,000 times to validate the resulting phylogenetic trees.

2. Results

2.1 Primary Structure and Physicochemical Properties of *C. canephora* WRKY Proteins

Using the ProtParam online tool (Zhou et al., 2018; Wang et al., 2019; Sha et al., 2019), 49 members of the *C. canephora* WRKY transcription factor gene family were identified. The primary structures and physicochemical properties of these proteins were compiled and analyzed. Results indicated that the total amino acid number of *C. canephora* WRKY gene family proteins ranged from 103 to 994 residues, with an average of 376 amino acids. Most were acidic proteins, with only 19 members having theoretical isoelectric points greater than 8.0. Instability coefficient analysis revealed that Cc02_{g05280} and Cc10_{g06400} were stable proteins, while the remainder were unstable (proteins with instability coefficients below 40 are considered stable). Two members (Cc01_{g14950} and Cc07_{g03120}) exhibited equal numbers of negatively and positively charged residues, displaying electrical neutrality. Twenty-five members had more negatively charged residues than positively charged residues (negatively charged), while 22 members had more positively charged residues (positively charged) (Table 1).

Comparative analysis of amino acid composition revealed that 44 members had serine as the most abundant amino acid, with content ranging from 7.5% to 16.6% (Figure 1 [Figure 1: see original paper]: a, b, c). Two members (Cc02_{g38600} and Cc03_{g00670}) had proline as the most abundant amino acid (Figure 1: d, e), one member (Cc05_{g04000}) had asparagine (Figure 1:

f), one member (Cc09_{g01430}) had alanine (Figure 1: g), and one member (Cc10_{g06400}) had lysine (Figure 1: h).

2.2 Prediction of Amino Acid Hydrophobicity/Hydrophilicity

Hydrophobicity/hydrophilicity prediction using the ProtScale online tool revealed that all transcription factor amino acid scores were negative, indicating they are hydrophilic proteins. The average hydrophilicity coefficients of the 49 WRKY proteins ranged from -1.196 to -0.473, confirming their hydrophilic nature. Maximum hydrophobicity values ranged from 0.289 to 2.589, with Cc07_{g03730} showing the highest and Cc08_{g11060} the lowest. Minimum hydrophobicity values ranged from -3.967 to -2.511, with Cc02_{g05280} showing the highest and Cc06_{g03470} the lowest. Cc08_{g11060} exhibited the smallest average hydrophilicity coefficient, indicating the strongest hydrophilicity, with maximum hydrophobicity at position 170 (value: 0.589) and minimum hydrophobicity at position 104 (value: -3.811) (Table 2).

2.3 Signal Peptide Prediction and Subcellular Localization

None of the 49 WRKY proteins contained signal peptides, suggesting they are non-secretory proteins. Subcellular localization prediction for all 49 WRKY proteins indicated that family members are localized to the nucleus (Table 1), demonstrating that CcWRKY family proteins function within the nuclear compartment.

2.4 Analysis of Encoded Amino Acid Transmembrane Structure

Transmembrane structure analysis of *C. canephora* WRKY proteins using the TMHMM Server (v.2.0) online program (Zheng et al., 2018) revealed that only one member, Cc02_{g22190}, possessed a distinct transmembrane region between amino acids 536 and 553 (Figure 2 [Figure 2: see original paper]). The remaining 48 members lacked transmembrane domains, indicating that *C. canephora* WRKY proteins are not transmembrane proteins.

2.5 Phosphorylation Site Analysis

Analysis revealed that serine phosphorylation sites were most abundant among the 49 WRKY proteins, followed by threonine and tyrosine. Serine showed the highest potential as a phosphorylation site, with minimum values reaching 0.995, substantially exceeding the threshold of 0.5. This suggests that CcWRKY gene family proteins may regulate their functions through phosphorylation at specific sites. Cc07_{g03730} contained the highest numbers of serine (83), threonine (39), and tyrosine (12) phosphorylation sites (Table 3).

2.6 Prediction and Analysis of Encoded Protein Secondary and Tertiary Structures

Protein secondary structure refers to regular, repeating conformations within polypeptide chains. Analysis of CcWRKY family members revealed that their secondary structures primarily consist of random coil, α-helix, β-turn, and extended strand—the four main secondary structural elements commonly investigated (Zhang et al., 2018). Random coil predominated, comprising 46.67% to 77.26% of secondary structures and primarily serving to connect other structural elements. β-turns represented the smallest proportion. The secondary structures were relatively uniform across the family, with all members except Cc00_{g06830}, Cc05_{g14660}, Cc07_{g16400}, Cc08_{g11060}, and Cc10_{g06400} following the pattern: random coil > α-helix > extended strand > β-turn. The exceptions showed the pattern: random coil > extended strand > α-helix > β-turn (Table 4 , Figure 3 [Figure 3: see original paper]).

Tertiary structure analysis revealed that *C. canephora* WRKY transcription factor family proteins could be classified into six categories (A, B, C, D, E, and F). Category D contained the most members (25), followed by category A (9 members), category E (7 members), category B (4 members), category C (3 members), and category F (1 member). Stability ranking from strongest to weakest was: F > D, E > A > C > B. Notably, Cc07_{g03730} matched the c5yvfD structure template with 315 amino acid residues at 100% confidence, while Cc05_{g08580} matched the c2aydA template with 75 residues at 100% confidence (Figure 4 [Figure 4: see original paper]).

2.7 Analysis of Amino Acid Sequence Domains

MEME online tool analysis of *C. canephora* WRKY gene family members identified six conserved motifs (Motif1–Motif6) based on conservation strength. Sequence logos revealed that the WRKYGQK heptapeptide motif was present in both Motif1 and Motif5. All 49 WRKY family members contained Motif1 and Motif2, suggesting Motif1 represents the WRKY domain. Each member possessed at least one WRKY heptapeptide, with maximum of five and minimum of two conserved motifs. Nineteen members contained only Motif1 and Motif2, while seven members contained Motif1, Motif2, Motif4, Motif5, and Motif6 (Figure 5 [Figure 5: see original paper], Table 5).

2.8 Construction of *C. canephora* WRKY Protein Phylogenetic Tree

Multiple sequence alignment of WRKY gene family member protein sequences and phylogenetic tree construction using MEGA software revealed six distinct groups (Group 1–Group 6). Group 1 contained the most members (14, including WRKY 12, 13, 23, 28, 48, 50, 51 [Cc07_{g16400} and Cc10_{g06400}], 56 [Cc02_{g38600} and Cc08_{g15290}], 57, 71, and 75 [Cc07_{g00980} and Cc11_{g17170}]), representing approximately 28.4% of the total. Groups 2, 5, and 3 contained 10, 10, and 9 members, respectively. Groups 6 and 4 were small-

est, with 5 and 1 member(s), respectively (Table 1). Phylogenetic relationships indicated that Groups 1–5 clustered together, suggesting close relationships, while Group 6 was more distantly related, implying that Group 6 members may be highly conserved throughout long-term evolutionary processes and represent the most ancestral-like class.

Group-specific alignment revealed diverse conserved amino acids across groups, yet all groups shared the WRKY conserved motif. Following Eulgem et al. (2000) classification criteria, the 49 WRKY gene family members were categorized into three families: Group 2 (D731-H785, D943-H998) belongs to WRKY Family I, containing two conserved heptapeptide structures (WRKYGQK) and two C2H2-type zinc finger structures; Group 6 (D151-C209) belongs to WRKY Family III, containing one conserved heptapeptide (WRKYGQK) and one C2HC-type zinc finger structure (Bade et al., 2017). Groups 1 (D205-H260), 3 (D359-H415), 4 (D151-H206), and 5 (D218-H274 or P316-H374) contain one conserved WRKY domain and one C2H2 zinc finger domain, classifying them as WRKY Family II (Figure 7 [Figure 7: see original paper]).

2.9 Phylogenetic Analysis of *C. canephora* WRKY Transcription Factors with Other Species

To further analyze the evolutionary relationships and patterns of *C. canephora* WRKY genes, WRKY1 protein sequences from 12 species including *Arabidopsis*, maize, and kiwifruit were selected for homologous clustering analysis. Phylogenetic tree construction using MEGA software revealed two major branches: tobacco, *C. canephora*, kiwifruit, *Arabidopsis*, Chinese chestnut, jujube, and apple clustered in Branch 1, while maize, rice, moso bamboo, banana, and oil palm clustered in Branch 2. Within Branch 1, *C. canephora* WRKY1 showed the closest relationship and highest sequence homology with tobacco, followed by kiwifruit, and was most distantly related to oil palm. From a dicotyledonous perspective, *C. canephora* clustered with kiwifruit, tobacco, *Arabidopsis*, Chinese chestnut, jujube, and apple, consistent with dicot classification. Monocotyledonous plants (maize, rice, moso bamboo, banana, and oil palm) formed a separate cluster, indicating a clear genetic divergence between dicots and monocots. From a woody plant perspective, *C. canephora* WRKY1 was most closely related to kiwifruit and most distant from oil palm. From an herbaceous perspective, it was closest to tobacco and most distant from rice (Figure 8 [Figure 8: see original paper]).

3. Discussion and Conclusion

WRKY transcription factors specifically recognize the W-box [(T)TGACC(AT)] motif in target gene promoters through their conserved amino acid sequence (WRKYGQK) to regulate transcriptional expression (Liang and Dong, 2018). Studies have successfully identified and characterized the biological features and molecular functions of AtWRKY and OsWRKY gene families (Rushton et al., 2010; Gu et al., 2015). WRKY gene families play essential roles

in seed germination, plant defense responses, metabolic regulation, and developmental/senescence control (Liang and Dong, 2018). In *Arabidopsis*, WRKY transcription factors can enhance tolerance to low phosphorus stress, with AtWRKY3 and AtWRKY4 positively regulating responses to necrotrophic pathogens (Lai et al., 2008; Chen et al., 2010). Tobacco studies have demonstrated WRKY involvement in leaf senescence and virus infection responses, with NtWRKY40 potentially participating in mechanical wounding and pathogen defense (Liu et al., 2016). Kiwifruit research indicates WRKY importance in waterlogging stress responses (Zhang et al., 2015). In jujube, *Candidatus Phytoplasma ziziphi* infection induces expression of ZjWRKY8, ZjWRKY52, ZjWRKY61, and ZjWRKY69 (Fu et al., 2018). Apple salt stress studies show that MdWRKY18 and MdWRKY40, induced by salt stress, can form homodimers or heterodimers to enhance salt tolerance (Xu et al., 2018). Overall, WRKY transcription factors play indispensable roles in plant life activities (Gao et al., 2016).

This study revealed that *C. canephora* WRKY family proteins contain 103–994 amino acids (average: 376), with isoelectric points ranging from 5.03 to 9.62 —similar to findings in jujube (Xue et al., 2018) and apple (Bhattarai, 2010). All *C. canephora* WRKY family proteins lack transmembrane domains, are hydrophilic and stable, contain no signal peptides, and are nuclear-localized, suggesting they function in the nucleus after cytoplasmic synthesis. The presence of phosphorylation sites at serine, threonine, and tyrosine residues suggests functional regulation via phosphorylation. Secondary structure prediction indicated random coil predominance, while tertiary structures were classified into six categories, with category D being most prevalent. Studies suggest AtWRKY61, AtWRKY72, AtWRKY6, and AtWRKY7 participate in plant immune signaling, defense responses, leaf senescence, and innate immunity (Robatzek et al., 2002; Bhattarai, 2010; Journot-Catalino et al., 2006). The stable D-class structure in *C. canephora* contains homologs to these *Arabidopsis* WRKYs, implying similar functional roles.

Conserved domain analysis revealed that five members (Cc00_{g01860}, Cc00_{g05270}, Cc00_{g05280}, Cc00_{g39120}, and Cc08_{g15910}) contain only two conserved motifs (Motif1 and Motif2), clustering them into one major group consistent with phylogenetic tree results (Group 6). The WRKY domain contains both the conserved heptapeptide (WRKYGQK) and a zinc finger structure (C2HC), which is characteristic of WRKY Family III (Li et al., 2019), consistent with findings in jujube and other species (Xue et al., 2018). Evolutionary analysis demonstrated that *C. canephora* WRKY transcription factor family proteins can be divided into six groups, with WRKY1 showing closest relationship to tobacco and most distant relationship to oil palm.

In conclusion, this study identified 49 WRKY genes in the *C. canephora* genome through bioinformatics approaches and conducted detailed analyses of their encoded proteins' physicochemical properties, phosphorylation sites, transmembrane structures, secondary and tertiary structures, functional domains, and

molecular evolution. These findings provide theoretical foundations and reference values for future research on *C. canephora* WRKY molecular functions, evolutionary origins, and stress response mechanisms (Bade et al., 2017).

References

BAKSHI M, OELMÜLLER R, 2014. WRKY transcription factors: jack of many trades in plants[J]. *Plant Signal Behav*, 9(2): e27700.

BAO CY, DENG L, ZHOU J, et al., 2018. Genome-wide identification and analysis of WRKY transcription factor family in kiwifruit[J]. *Mol Plant Breed*, 16(14): 4473-4488.

BHATTARAI KK, ATAMIAN HS, KALOSHIAN I, et al., 2010. WRKY72-type transcription factors contribute to basal immunity in tomato and *Arabidopsis* as well as gene-for-gene resistance mediated by the tomato gene Mi-1[J]. *Plant J*, 63(2): 229-240.

BUDE RG, BAO ML, ZHANG HX, et al., 2017. Bioinformatics analysis of WRKY transcription factor genes family in melon[J]. *Genom Appl Biol*, 36(11): 4761-4769.

CHEN YF, XU Q, WANG H, et al., 2010. Functional analysis of *Arabidopsis* WRKY transcription factor in plant response to low phosphorus stress[G]. Abstract Compilation of Papers on Plant Cell Biology 2010 Annual Meeting of the Chinese Botany Society.

CHENG T, WEI Q, LI GL, 2016. Bioinformatics analysis of the TPS gene family in *Coffea canephora*[J]. *Chin Bull Bot*, 51(2): 235-250.

DAVIS AP, TOSH J, RUCH N, et al., 2011. Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*[J]. *Bot J Lin Soc*, 167(4): 357-377.

DING M, CHEN J, JIANG Y, et al., 2015. Genome-wide investigation and transcriptome analysis of the WRKY gene family in *Gossypium*[J]. *Mol Genet Genomics*, 290(1): 151-171.

DONG X, YANG Y, ZHANG Z, et al., 2019. Genome-wide identification of WRKY genes and their response to cold stress in *Coffea canephora*[J]. *Forests*, 10(4): 335.

EULGEM T, RUSHTON PJ, ROBATZEK S, et al., 2000. The WRKY superfamily of plant transcription factors[J]. *Trends Plant Sci*, 5: 200-206.

FU B, YE X, WANG HY, et al., 2018. Identification of WRKY transcription factors in jujube and their responses to '*Candidatus* Phytoplasma ziziphi' and salicylic acid or methyl jasmonate treatments[J]. *Sci Silva Sin*, 54(8): 65-78.

GAO R, LIU P, YONG YH, et al., 2016. Genome-wide transcriptomic analysis reveals correlation between higher WRKY61 expression and reduced symptom severity in turnip crinkle virus infected *Arabidopsis thaliana*[J]. *Sci Rep*, 6(1): 24604.

GU YB, JI ZR, CHI FM, et al., 2015. Bioinformatics and expression analysis of the WRKY gene family in apple[J]. *Sci Agric Sin*, 48(16): 3221-3238.

ISHIGURO S, NAKAMURA K, 1994. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and β-amylase from sweet potato[J]. *Mol Gen Genet*, 244(6): 563-571.

JIANG Y, DEYHOLOS MK, 2006. Comprehensive transcriptional profiling of NaCl-stressed *Arabidopsis* roots reveals novel classes of responsive genes[J]. *BMC Plant Biol*, 6(1): 25.

JOURNOT-CATALINO N, SOMSSICH IE, ROBY D, et al., 2006. The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana*[J]. *Plant Cell*, 18(11): 3289-3302.

KELLEY LA, MEZULIS S, YATES CM, et al., 2015. The Phyre2 web portal for protein modeling, prediction and analysis[J]. *Nat Protoc*, 10(6): 845-858.

KUMAR S, STECHER G, LI M, et al., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms[J]. *Mol Bio Evol*, 35(6): 1547-1549.

LAGACÉ M, MATTON DP, 2004. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*[J]. *Planta*, 219(1): 185-189.

LAI Z, VINOD KM, ZHENG Z, et al., 2008. Roles of *Arabidopsis* WRKY3 and WRKY4 transcription factors in plant responses to pathogens[J]. *BMC Plant Biol*, 8(1): 68.

LI XY, XU HX, CHEN JW, 2019. Identification and expression analysis of WRKY transcription factors in *Eriobotrya japonica*[J]. *Acta Hort Sin*, 46(5): 939-954.

LI Z, HAN X, ZHANG YH, et al., 2019. Cloning and functional analysis of a phosphorus stress response gene *GhWRKY6* in cotton (*Gossypium hirsutum* L.)[J]. *Cotton Sci*, 31(3): 182-191.

LIANG B, DONG D, 2018. Advances in WRKY transcription factors in plants[J]. *Bull Biol*, 53(6): 5-8.

LIU JJ, CHENG CL, XI YZ, et al., 2016. Roles of *Nicotiana tobacum* NtWRKY40 in plant responding to virus infection[J]. *Biotechnol Bull*, 32(10): 188-198.

MONENTE C, LUDWIG IA, IRIGOYEN A, et al., 2015. Assessment of total (free and bound) phenolic compounds in spent coffee extracts[J]. *J Agric Food*

Chem, 63(17): 4327-4334.

ROBATZEK S, SOMSSICH IE, 2010. A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence- and defence-related processes[J]. *Plant J*, 28(2): 123-133.

ROBATZEK S, SOMSSICH IE, 2002. Targets of AtWRKY6 regulation during plant senescence and pathogen defense[J]. *Gene Dev*, 16(9): 1139-1149.

RODRIGUES NP, SALVA TDJ, BRAGAGNOLO N, 2015. Influence of coffee genotype on bioactive compounds and the *in vitro* capacity to scavenge reactive oxygen and nitrogen species[J]. *J Agric Food Chem*, 63(19): 4815-4826.

RUSHTON DL, TRIPATHI P, RABARA RC, et al., 2012. WRKY transcription factors: key components in abscisic acid signalling[J]. *Plant Biotechnol J*, 10(1): 2-11.

RUSHTON PJ, SOMSSICH IE, RINGLER P, et al., 2010. WRKY transcription factors[J]. *Trends Plant Sci*, 15(5): 247-258.

SAITOU N, NEI M, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees[J]. *Mol Biol Evol*, 4(4): 406-425.

SHA T, LI YJ, ZHANG FB, et al., 2019. Antigenic epitopes of *Echinococcus granulosus* proteins: egG1y162-1 and egG1y162-2: a comparative study[J]. *J Xinjiang Med Univ*, 42(1): 12-19.

SU L, WANG PF, YANG Y, et al., 2019. Genome-wide identification and analysis of WRKY transcription factors in grape[J]. *Heilongjiang Agric Sci*, (1): 13-22.

WANG PY, QU SS, LIANG Y, et al., 2019. Bioinformatics analysis of WRKY transcription factor family in rice[J]. *J Anhui Agr Sci*, 47(12): 123-126.

WU X, SHIROTO Y, KISHITANI S, et al., 2009. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing *OsWRKY11* under the control of *HSP101* promoter[J]. *Plant Cell Rep*, 28(1): 21-30.

XIE ZW, WANG LJ, CHEN JY, et al., 2016. Studies on WRKY transcription factors and their biological functions in plants[J]. *J Agric Sci Technol China*, 18(3): 46-54.

XU HF, YANG GX, ZHANG J, et al., 2018. Molecular mechanism of apple MdWRKY18 and MdWRKY40 participating in salt stress[J]. *Sci Agric Sin*, 51(23): 4514-4521.

XU YH, WANG JW, WANG S, et al., 2004. Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-δ-cadinene synthase-A1[J]. *Plant Physiol*, 135(1): 507-515.

XUE BP, ZHANG H, ZU HH, et al., 2018. Bioinformatics analysis of the WRKY transcription factor family in *Ziziphus jujuba*[J]. *Mol Plant Breed*, 16(14): 4543-4558.

YAN L, HUANG LF, WANG XY, et al., 2018. Comparative trial of robusta coffee cultivars[J]. *Chin J Trop Crop*, 39(7): 1276-1281.

YANG HG, 2014. Effects of drought stress on lipid peroxidation, antioxidant enzyme activities and osmolyte content in leaves of *Coffea canephora* seedlings[J]. *Chin J Trop Crop*, 35(5): 944-949.

YU CS, CHEN YC, LU CH, et al., 2006. Prediction of protein subcellular localization[J]. *Proteins*, 64(3): 643-651.

ZHANG CL, MAN LL, XIANG DJ, et al., 2018. Cloning and bioinformatics analysis of *BjSnRK2C* gene from potherb mustard[J]. *Acta Agric Boreal-Sin*, 33(6): 24-32.

ZHANG H, 2017. Bioinformatics analysis of WRKY transcription factor family and identification of some genes related to stress resistance in tomato[D]. Harbin: Northeast Agricultural University.

ZHANG JY, HUANG SN, MO ZH, et al., 2015. Transcriptional sequencing and comparative analysis of the different expression genes of *Actinidia chinensis* under drought stress[C]. Annual Meeting of Chinese Society for Horticultural Science in 2015.

ZHANG L, HE XJ, LI C, et al., 2019. Expression and functional prediction of WRKY transcription factors in leaf senescence of tobacco 'Yunyan 87' [J]. *Mol Plant Breed*, 17(8).

ZHENG C, ZHENG ES, WANG YM, et al., 2018. Research progress on rice WRKY transcription factors[J]. *Letters in Biotechnol*, 29(2): 286-294.

ZHOU J, HUANG J, 2018. Cloning and functional identification of chitinase gene *SlChi* in *Salix*[J]. *Mol Plant Breed*, 16(24): 8013-8021.

ZOU Z, 2013. Genome-wide identification and phylogenetic analysis of WRKY transcription factor family in castor bean[J]. *Oil Crop Sci*, 35(1): 36-42.

Note: Figure translations are in progress. See original paper for figures.