

RESEARCH ARTICLE

Genome-wide identification and characterization of cacao WRKY transcription factors and analysis of their expression in response to witches' broom disease

Dayanne Silva Monteiro de Almeida¹, Daniel Oliveira Jordão do Amaral¹, Luiz-Eduardo Del-Bem², Emily Bronze dos Santos¹, Raner José Santana Silva¹, Karina Peres Gramacho³, Michel Vincentz², Fabienne Micheli^{1,4*}

1 Universidade Estadual de Santa Cruz (UESC), Departamento de Ciências Biológicas (DCB), Centro de Biotecnologia e Genética (CBG), Rodovia Ilhéus-Itabuna, km 16, Ilhéus-BA, Brazil, **2** Centro de Biologia Molecular e Engenharia Genética, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brasil, **3** Cocoa Research Center, CEPLAC/CEPEC, Itabuna-BA, Brazil, **4** CIRAD, UMR AGAP, Montpellier, France

* fabienne.micheli@cirad.fr



OPEN ACCESS

Citation: Silva Monteiro de Almeida D, Oliveira Jordão do Amaral D, Del-Bem L-E, Bronze dos Santos E, Santana Silva RJ, Peres Gramacho K, et al. (2017) Genome-wide identification and characterization of cacao WRKY transcription factors and analysis of their expression in response to witches' broom disease. PLoS ONE 12(10): e0187346. <https://doi.org/10.1371/journal.pone.0187346>

Editor: Martina Stromvik, McGill University, CANADA

Received: April 10, 2017

Accepted: October 18, 2017

Published: October 30, 2017

Copyright: © 2017 Silva Monteiro de Almeida et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research was supported by FAPESB (project DTE0038/2013) coordinated by FM. DSMA took a practical course at Unicamp funded by Coordenação de Aperfeiçoamento Pessoal de Nível

Abstract

Transcriptional regulation, led by transcription factors (TFs) such as those of the WRKY family, is a mechanism used by the organism to enhance or repress gene expression in response to stimuli. Here, we report on the genome-wide analysis of the *Theobroma cacao* WRKY TF family and also investigate the expression of WRKY genes in cacao infected by the fungus *Moniliophthora perniciosa*. In the cacao genome, 61 non-redundant WRKY sequences were found and classified in three groups (I to III) according to the WRKY and zinc-finger motif types. The 61 putative WRKY sequences were distributed on the 10 cacao chromosomes and 24 of them came from duplication events. The sequences were phylogenetically organized according to the general WRKY groups. The phylogenetic analysis revealed that subgroups IIa and IIb are sister groups and share a common ancestor, as well as subgroups IIc and IIe. The most divergent groups according to the plant origin were IIc and III. According to the phylogenetic analysis, 7 *TcWRKY* genes were selected and analyzed by RT-qPCR in susceptible and resistant cacao plants infected (or not) with *M. perniciosa*. Some *TcWRKY* genes presented interesting responses to *M. perniciosa* such as Tc01_p014750/Tc06_p013130/AtWRKY28, Tc09_p001530/Tc06_p004420/AtWRKY40, Tc04_p016130/AtWRKY54 and Tc10_p016570/AtWRKY70. Our results can help to select appropriate candidate genes for further characterization in cacao or in other *Theobroma* species.

Introduction

Plants, whether growing under natural or agricultural conditions, are exposed to adverse environmental situations that affect their development and can drastically reduce their productivity. Such environmental stimuli can be abiotic (e.g., drought, cold, wounds) or caused by

Superior (CAPES, PROCAD project). DSMA and RJSS were funded by FAPESB. DOJJA was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). FM and KPG received a Productivity Grant from CNPq (PQ1). LEDB was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). EBS was funded by UESC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ABA, abscisic acid; COC, calcium oxalate crystals; hai, hours after inoculation; dai, days after inoculation; JA, jasmonic acid; LRR, leucine-rich repeat; OA, oxalic acid; ROS, reactive oxygen species; SA, salicylic acid; TF, transcription factor; WD, WRKY domain.

pathogens [1]. However, plants have mechanisms to survive in such adverse conditions, including tolerance or resistance to stress through adaptation mechanisms [2]. When stress conditions are detected by the plants, complex transduction pathways are induced, initiating a series of molecular, physiological and metabolic events, generally leading to an increase of tolerance/resistance [3]. Transcriptional regulation, led by transcription factors (TFs)±regulation proteins that link DNA sequences in specific promoter regions of target genes±is the first mechanism that is activated by the organism to enhance or repress gene expression involved in response to internal or external stimuli [4]. Such TFs can regulate more than one gene as well as other TFs [4].

Among TFs, the WRKY family is well known. The first WRKY protein was identified in sweet potato [5] and since then, WRKY proteins have been characterized in several other plants [6, 7] as well as in algae and non-plant eukaryotes [8]. The WRKY TFs bind a specific promoter sequence of the target gene, known as a W-box, positively or negatively regulating the target gene expression. The WRKY proteins contain one or two DNA binding domains of 60 amino acids containing the conserved heptapeptide WRKYGQK followed by a zinc-finger motif C₂H₂ (C-X₄₋₅-C-X₂₂₋₂₃-H-X-H) or C₂HC (C-X₇-C-X₂₃₋₂₄-H-X-C) [6]. In some species, members of the WRKY family containing three DNA binding domains have been found [9, 10]. The WRKY TF family is known to be involved in response to biotic and abiotic stresses [7], and also to modulate various other processes in plants such as embryogenesis [11], trichome and seed development [12], leaf senescence [13], fruit and pollen development [14], biomass [15], secondary metabolite biosynthesis [16] and hormone signalling [17]. In *Arabidopsis* species, several WRKY genes have been experimentally characterized and associated with response to fungal or bacterial pathogens [18, 19], as well as to nematodes [20]. In cacao (*Theobroma cacao* L.), a portion of WRKY genes have been previously identified and analyzed with regards to their expression, phylogeny within the Malvaceae family, and/or for their potential use in marker assisted selection [21±23], but no exhaustive analysis of this TF family based on cacao genome data has been carried out.

The cacao tree is cultivated mainly for its beans, used as raw material for making chocolate. However, bean production is threatened worldwide by several pathogens, such as *M. perniciosa*, the agent of the witches' broom disease [24], one of the most devastating diseases of the crop in Central and South America, and the Caribbean [25]. This disease has been responsible for the abandonment by producers of many cultivated areas in these regions [26], to the point of causing a shortage of cacao beans in the global market [27]. In this scenario, during the genomic and post-genomic eras, tools have been developed to understand the cacao-*M. perniciosa* interaction as well as to identify molecules that can be used to develop disease control methods. Here, we report on the genome-wide analysis of the *T. cacao* WRKY TF family and the identification of a comprehensive and non-redundant set of WRKY genes from this species. Subsequently, chromosomal location was determined and phylogenetic and motif analyses performed as a base for further comparative genomics studies. Moreover, expression patterns of WRKY genes in cacao infected (or not) with the pathogenic fungus *M. perniciosa* were also investigated. From the 61 TcWRKY proteins identified some were potentially involved cacao's response to *M. perniciosa* and can be considered good candidates for subsequent functional analyses or disease management.

Material and methods

Datasets and WRKY protein identification

The scheme of the *in silico* pipeline used for identification of the *Theobroma cacao* WRKY protein set is shown in the S1 Fig. The *Theobroma cacao* protein sequences were downloaded

from the CacaoGenDB database v1.0 (*Theobroma_cacao_v1.pep.faa.gz*; <http://cocoa-genome-hub.southgreen.fr/gbrowse>) [28]. The WRKY protein sequences of *Arabidopsis thaliana* were downloaded from the Phytozome database v12.1 (www.phytozome.org) and plant transcription factor database v.3.0 (<http://plntfdb.bio.uni-potsdam.de/v3.0>); only the non-redundant proteins between both databases were considered and used for the subsequent analysis (S1 Fig). The cacao proteins were screened to search for WRKY members using the local BLASTP program (blastall version 2.2.27; <ftp://ftp.ncbi.nih.gov/blast/executables/blast+/LATEST>) and the WRKY protein sequences of *A. thaliana* as input sequences (S1 Fig). The e-value for BLASTP was set at 1e-10 to obtain the final dataset of WRKY proteins. When more than one alternative splicing sequence was found for the same locus, only the longest non-redundant sequences were used for subsequent analyses. Subsequently, these candidate TcWRKY protein sequences were submitted to analysis using the InterPro (<http://www.ebi.ac.uk/interpro/>) and PFAM programs (<http://pfam.xfam.org/>) to confirm the presence of the WRKY domain, and analyzed using the SMART program (<http://smart.embl-heidelberg.de/>) to confirm the presence of the zinc-finger domain (S1 Fig). The protein sequences lacking both the WRKY and the zinc-finger domains were manually excluded. In parallel, an automatic search of TcWRKY proteins was made using the browser tool of CacaoGenDB v.1.0 based on keywords and InterPro numbers. The comparison of results from both analyses (local BlastP vs automatic search) allowed a fine, precise and complete analysis guaranteeing the identification of the largest non-redundant TcWRKY TF set (S1 Fig). This final largest non-redundant TcWRKY TF set contained 61 proteins (Table 1). Two of the TcWRKY proteins presented a highly altered WRKY motif and/or zinc-finger motif (Tc00_g017240 and Tc02_g001170) and for this reason were excluded from the phylogenetic analysis (Table 1). Moreover, one of the TcWRKY proteins presented the zinc-finger motif but not the conserved heptapeptide (Tc02_g012180) for this reason was excluded from all analyses (Table 1). Thus, depending on the analysis, only 60 or 58 protein sequences were used.

TcWRKY gene classification and chromosomal location

After confirmation and identification of the final dataset of TcWRKY proteins (S1 Fig), all whole protein sequences, except the Tc02_g012180, were used for sequence alignment using the ClustalOMEGA software (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). The cacao TcWRKY proteins were categorized based on *Arabidopsis* WRKY protein classification [6]. The distribution of the TcWRKY sequences on cacao chromosomes was obtained from the CacaoGenDB database v1.0 (<http://cocoagendb.cirad.fr>) [28] using "WRKY" as input in the "Search by Keywords" tool (<http://cocoa-genome-hub.southgreen.fr/content/search-keywords>). The sequences downloaded from the CacaoGenDB database v1.0, excluding the chromosome Tc00, were used as an input file for prediction of WRKY gene duplication and collinearity using the MCScanX toolkit, according to the manual [29]. This analysis was made independently from other previously published reports [28] mainly because the CacaoGenDB database was regularly updated since the database creation and work publication.

Identification of conserved motifs in TcWRKY proteins

The detection of the motif composition in the 58 identified cacao WRKY proteins (all proteins except Tc00_g017240, Tc02_g001170 and Tc02_g012180) was performed with the MEME 4.9.1 program (<http://meme.nbcr.net/meme/intro.html>) [30]. MEME represents motifs as position-dependent letter-probability matrices which describe the probability of each possible letter at each position in the pattern [30]. The maximum number of motifs was set at 20, the maximum motif length was set at 80 amino acids, the optimum motif width was constrained

Table 1. WRKY proteins present in the cacao genome (CacaoGenDB v1.0 [28]). The variants of the conserved WRKYGQK peptide are shown in bold. nd: not determined.

Gene locus	WRKY domain			Group	Chromosome	Accession number
	Conserved heptapeptide	Zinc-finger type	Domain amount			
Tc02_g032670	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	2	XP_007045981.1
Tc04_g009710	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	4	XP_007032626.1
Tc09_g034740	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	9	XP_007017012.1
Tc05_g001480	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	5	XP_007026547.1
Tc05_g005710	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	5	XP_007027196.1
Tc05_g020810	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	5	XP_007029376.1
Tc07_g002020	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	7	XP_007020620.1
Tc07_g000190	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	7	XP_007020301.1
Tc01_g018460	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	1	XP_007049282.1
Tc09_g002780	WRKYGQK/WRKYGQK	C ₂ H ₂	2	I	9	XP_007011940.1
Tc09_g001530	WRKYGQK	C ₂ H ₂	1	Ila	9	XP_007011727.1
Tc09_g001520	WRKYGQK	C ₂ H ₂	1	Ila	9	XP_007011726.1
Tc06_g004420	WRKYGQK	C ₂ H ₂	1	Ila	6	XP_007023430.1
Tc06_g019530	WRKYGQK	C ₂ H ₂	1	Ilb	6	XP_007026134.1
Tc02_g003250	WRKYGQK	C ₂ H ₂	1	Ilb	2	XP_007041570.1
Tc01_g032940	WRKYGQK	C ₂ H ₂	1	Ilb	1	XP_007051308.1
Tc01_g017430	WRKYGQK	C ₂ H ₂	1	Ilb	1	XP_007049086.1
Tc03_g009820	WRKYGQK	C ₂ H ₂	1	Ilb	3	XP_007037253.1
Tc07_g002910	WRKYGQK	C ₂ H ₂	1	Ilb	7	XP_007020766.1
Tc04_g007790	WRKYGQK	C ₂ H ₂	1	Ilb	4	XP_007032468.1
Tc02_g033950	WRKYGQK	C ₂ H ₂	1	Ilb	2	XP_007046206.1
Tc04_g029800	WRKYGQK	C ₂ H ₂	1	Ilc	4	XP_007035842.1
Tc01_g010370	WRKYGQK	C ₂ H ₂	1	Ilc	1	XP_007048165.1
Tc06_g013130	WRKYGQK	C ₂ H ₂	1	Ilc	6	XP_007025165.1
Tc01_g039500	WRKYGQK	C ₂ H ₂	1	Ilc	1	XP_007052352.1
Tc01_g014750	WRKYGQK	C ₂ H ₂	1	Ilc	1	XP_007048873.1
Tc00_g047270	WRKYGQK	C ₂ H ₂	1	Ilc	0	XP_007043929.1
Tc04_g004210	WRKYGQK	C ₂ H ₂	1	Ilc	4	XP_007031915.1
Tc01_g031960	WRKYGQK	C ₂ H ₂	1	Ilc	1	XP_007051149.1
Tc05_g004380	WRKYGQK	C ₂ H ₂	1	Ilc	5	XP_007026988.1
Tc00_g076580	WRKYGKK	C ₂ H ₂	1	Ilc	0	XP_007037699.1
Tc01_g035290	WRKYGQK	C ₂ H ₂	1	Ilc	1	XP_007051697.1
Tc08_g013540	WRKYGKK	C ₂ H ₂	1	Ilc	8	XP_007019683.1
Tc03_g015140	WRKYGQK	C ₂ H ₂	1	Ilc	3	XP_007038026.1
Tc01_g010220	WRKYGKK	C ₂ H ₂	1	Ilc	1	XP_007048143.1
Tc00_g017270	WRKYGKK	C ₂ H ₂	1	Ilc	0	XP_007037701.1
Tc09_g005290	WRKYGKK	C ₂ H ₂	1	Ilc	9	XP_007012369.1
Tc02_g032350	WRKYGQK	C ₂ H ₂	1	Ilc	2	XP_007045916.1
Tc05_g027100	WRKYGQK	C ₂ H ₂	1	Ild	5	XP_007030538.1
Tc09_g000780	WRKYGQK	C ₂ H ₂	1	Ild	9	XP_007011614.1
Tc01_g005580	WRKYGQK	C ₂ H ₂	1	Ild	1	XP_007047364.1
Tc03_g025390	WRKYGQK	C ₂ H ₂	1	Ild	3	XP_007039903.1
Tc01_g027130	WRKYGQK	C ₂ H ₂	1	Ild	1	XP_007050397.1
Tc08_g000030	WRKYGQK	C ₂ H ₂	1	Ild	8	XP_007017155.1
Tc03_g028030	WRKYGQK	C ₂ H ₂	1	Ile	3	XP_007040356.1
Tc06_g000970	WRKYGQK	C ₂ H ₂	1	Ile	6	XP_007022995.1
Tc03_g019750	WRKYGQK	C ₂ H ₂	1	Ile	3	XP_007099721.1

(Continued)

Table 1. (Continued)

Gene locus	WRKY domain			Group	Chromosome	Accession number
	Conserved heptapeptide	Zinc-finger type	Domain amount			
Tc06_g013990	WRKYGQK	C ₂ H ₂	1	Ile	6	XP_007025322.1
Tc01_g035330	WRKYGQK	C ₂ H ₂	1	Ile	1	XP_007051705.1
Tc01_g031780	WRKYGQK	C ₂ H ₂	1	Ile	1	XP_007051114.1
Tc10_g016560	WRKYGQK	C ₂ HC	1	III	10	XP_007011366.1
Tc03_g017550	WRKYGQK	C ₂ HC	1	III	3	XP_007038956.1
Tc03_g028700	WRKYGQK	C ₂ HC	1	III	3	XP_007040478.1
Tc01_g034680	WRKYGQK	C ₂ HC	1	III	1	XP_007051596.1
Tc04_g016130	WRKYGQK	C ₂ HC	1	III	4	XP_007033512.1
Tc10_g016570	WRKYGQK	C ₂ HC	1	III	10	XP_007011367.1
Tc02_g001230	WRKHGQT	C ₂ HC	1	III	2	XP_007041160.1
Tc02_g001200	WRKYGQK	C ₂ HC	1	III	2	XP_007041155.1
Tc00_g017240 ^{a*}	WRCIGIK	C ₁ H ₂	1	nd	0	XP_007037703.1
Tc02_g001170 ^{a*}	RTKYYRC	C ₁ HC	1	nd	2	XP_007041153.1
Tc02_g012180 ^{a**}	No conserved stretch	C ₂ H ₂ ^b	1	nd	2	XP_007043029.1

^a Sequences not classified in any group due to the absence of WRKY domain, incomplete zinc finger motif and/or small conserved WRKY domain.

^b Zinc finger C₂H₂ identified by the Smart program (<http://smart.embl-heidelberg.de/>)

* Excluded from the phylogenetic analyses.

** Excluded from all the subsequent analyses.

<https://doi.org/10.1371/journal.pone.0187346.t001>

to be between 6 and 300 residues, and the other parameters were used with the default settings. The organization of the different motifs (e.g. WRKY DNA-binding domain, leucine rich repeat/LRR) present in the cacao WRKY proteins was performed with the SuperFamily database v.1.75 (<http://supfam.org/SUPERFAMILY/>; [31]) based on hidden Markov models.

Phylogenetic analysis

A phylogenetic tree of amino acid sequences of WRKY domains from cacao (58 sequences, see paragraph above) and *Arabidopsis* (S1 Table) was constructed. The amino acid sequences of WRKY domains were aligned using the MUSCLE program v3.6 [32] with default parameters. The MEGA 5.1 software was used to construct a rooted phylogenetic tree [33]. The tree based on WRKY domains of *Arabidopsis* and cacao was used to identify and classify putative orthologs. The statistical method used to construct the tree was neighbor-joining [34], the evolutionary distances were obtained using the p-distances method, and these distances were used to estimate the number of amino acid substitutions per site. The reliability of each tree was established by conducting 1000 bootstrap sampling steps. To construct the tree with all species used in this study, the JTT evolutionary model plus gamma-distributed rate (JTT+G) was used as determined by the Modeltest program version 3.7 [35]. The phylogenetic analysis was completed by an analysis of orthology using the plant transcription factor database v.3.0 (<http://plntfdb.bio.uni-potsdam.de/v3.0>).

Plant material

The plant material used in this study consisted in TSH1188 and Catongo *T. cacao* genotypes. TSH1188 was chosen based on its demonstrated resistance to witches' broom disease from field progeny trials assessed by the number of vegetative and cushion brooms per plant and

per year, and by witches' broom incidence scale [36, 37]. The susceptible cultivar Catongo was chosen as the standard for susceptibility. These genotypes were previously used as resistance and susceptibility standards in several molecular and histological studies of witches' broom disease [38–42]. Seedlings, derived from open-pollinated pods of both genotypes were planted in a mixture of commercial potting mix (Plantmax[®], Eucatex, São Paulo, SP, Brazil) and clay-rich soil, in a 2:1 proportion, and grown in sterile substrate in the greenhouses of CEPLAC/CEPEC (Bahia, Brazil) under natural light and 90% relative humidity until the inoculation day. All the experiments followed a complete randomized design.

Plant inoculation procedure

The inoculation procedure was conducted in the greenhouses of CEPLAC/CEPEC (Bahia, Brazil) using inoculum from *Moniliophthora perniciosa* isolate 4145, which has been maintained in the CEPLAC/CEPEC phytopathological *M. perniciosa* collection (CEGEN n^o109/2013/SECEX-GEN) in sterile distilled water [43] and in mineral oil. The inoculation procedure has been previously described in detail [42]. Briefly, apical shoot apices of 300 4-week-old seedlings from each genotype were inoculated with a 20 μ l drop of a basidiospore suspension (2.10^5 basidiospores ml^{-1} with >80% germination) in 0.3% agar [44]. Inoculated seedlings were incubated for 48 h in a control dark moist chamber at $23\pm 2^\circ\text{C}$ and relative humidity greater than 97%. Afterwards seedlings were transferred to an acclimatization greenhouse with $23\pm 2^\circ\text{C}$ temperature and irrigation for 20 minutes three times a day until the end of the experiment. The greenhouse relative humidity was around 80% controlled through an automated fogging system. Seedlings mock-inoculated with sterile 0.3% were used as controls. Symptoms were observed weekly up to 60 days after inoculation (dai). The inoculation efficiency was checked based on the inoculum viability (>80% of spore germination observed 24 h after inoculation [hai]) and on disease incidence on the susceptible cultivar Catongo (>80% of disease incidence observed 60 dai). Apical shoot apices were harvested at 6, 12, 24, 48 and 72 hai and 7, 15, 30 and 45 dai. These harvesting points correspond to the main disease stages as previously described [40–42, 45, 46]. Briefly, 6h corresponded to *M. perniciosa* penetration in resistant and susceptible genotypes. The 24 to 72 hai period corresponded to the early stages of the infection; at 48 h the fungus hyphae were observed in the cortex beneath the epidermal layer in the susceptible genotype (in the resistant genotype the infection and fungus progression were reduced or stopped). From 15 to 25 dai, macroscopically the symptoms on susceptible genotype were observed as apical swellings and slight morphological alterations of the shoots (S2 Fig), and microscopically, hyphal strands were seen in the cortex and grew toward the vascular bundles through xylem ray cells. At 30 dai, apical hypertrophy and swellings were observed (S2 Fig); at this time the pathogen reached the pith, and swollen, flexuous, septated and unclamped hyphae grew in the apoplast. The infection continued to develop forming the phenotypic response of terminal green brooms that could be seen at 45 dai (S2 Fig). At 60 dai, the infected plant presented macroscopic symptoms called dry broom (data not shown). Control plants were kept and harvested under the same conditions and at the same time points. For each genotype and at each harvesting time (for inoculated and control plants), 24 samples were collected (1 sample = 1 apical apex of 1 cacao plantlet) and immediately frozen in liquid nitrogen and stored at -80°C until use. Then three samples collected from each genotype at each harvesting time were pooled forming one biological replicate; two biological replicates were obtained (i.e., 6 apices from the 24 collected were used).

Total RNA extraction and cDNA synthesis

Cacao samples were macerated in liquid nitrogen until a fine powder was obtained. Total RNA was extracted from 100 ± 150 mg of macerated tissue using the RNeasy[®] Total RNA

isolation kit according to the manufacturer's instructions (Thermo Scientific) with modifications as previously described [47]. Briefly, after the addition of the lysis buffer to the macerated samples, a sonication step was added (10 s pulse/min, 70% output; Gex Ultrasonic processor 130, 130 W) to break polysaccharides which are present in high levels in cacao tissues. This step was conducted on ice. RNA was quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific) and its integrity was checked by 1% agarose gel electrophoresis. RNA was treated by DNase I RNase-free according to the manufacturer's instructions (Invitrogen). The cDNA was synthesized from 200 ng of RNA using the RevertAid First Strand cDNA kit according to the manufacturer's instruction (Thermo Scientific). The cDNA quantification was carried out in the same NanoDrop 2000 spectrophotometer.

Primer design and qPCR analysis

Seven cacao *WRKY* genes (Tc04_t016130, Tc10_t016570, Tc09_t001530, Tc06_t004420, Tc06_t013130, Tc01_t014750, Tc01_t018460) were selected based on the phylogenetic analysis and searches for genes that are well characterized and possibly involved in plant defense mechanisms against pathogenic fungi in the *Arabidopsis* genus. Specific primers were designed for each gene using the OligoPerfect™ Designer tool (<http://tools.thermofisher.com>) according to the following criteria: i) amplicon size of 65±150 bp; ii) primer length of 17±23 bases; iii) melting temperature of 57±63°C; and iv) GC content of 40%-80% (S2 Table and S3 Table). The OligoAnalyzer v.3.1 program (<https://www.idtdna.com/calc/analyzer>) was used to analyze the primer pairs in relation to hairpin loop, self-dimer and hetero-dimer formation (<https://www.idtdna.com/calc/analyzer>). Primers were also designed to amplify specific regions presenting different sizes, melting temperatures, GC contents and GC/AT ratios (S3 Table) to avoid cross-reaction between genes from the cacao *WRKY* family [48]. For qPCR analysis, two reference genes (malate dehydrogenase/MDH and glyceraldehyde 3-phosphate dehydrogenase/GAPDH) previously described in cacao meristems infected by *M. perniciosa* [47, 49] were used (S2 Table). Expression analysis by qPCR was conducted in an Agilent Technologies Stratagene Mx3005P system (Agilent Technologies). The qPCR reaction consisted of 200 ng of cDNA, 0.5 μM of each primer from candidate or reference genes (S2 Table) and 1X of Maxima™ SYBR Green/ROX qPCR Master Mix (Thermo Scientific) in a total volume of 12.5 μl. Cycling conditions were: 50°C for 2 min, 95°C for 1 min followed by 40 cycles at 95°C for 30 s, 58°C for 45 s and 72°C for 30 s, with detection of the fluorescent signal at the end of each extension cycle. To verify that each primer pair produced only a single PCR product, dissociation analysis was carried out under the following cycling conditions: 95°C for 25 s, 58°C for 30 s and 72°C for 30 s. The amplification efficiency of each primer pair was analyzed using three amounts (50, 100 and 200 ng) of each cDNA sample. Experiments also included a negative control (no template DNA). Real-time data acquisition was performed with the Stratagene MX3005P system containing the MxPro QPCR software (Agilent Technologies), which provided the values of cycle threshold (Ct) and of fluorescence. Amplification efficiency (E) was accessed using the Miner 2.2 software [50]. The gene expression level was analyzed with three experimental repetitions for both Catongo and TSH1188 genotypes with the comparative Ct method ($2^{-\Delta\Delta Ct}$) using: i) MDH and GAPDH as reference genes (average of expression values from both genes); and ii) non-inoculated plants as a calibrator (at each harvesting time, a non-inoculated sample was collected and used as calibrator of the corresponding inoculated sample). Statistical analysis was done using the SASM-Agri software v.8.2 [51], which tested the experiments as a completely randomized design. *t*-test and *F*-test (ANOVA) were applied with a critical value of 0.05. The Duncan test ($P \leq 0.05$) was employed for mean separation when *F*-values were significant.

Results

Identification and classification of TcWRKY sequences

In the cacao genome, 61 non-redundant sequences corresponding to putative WRKY proteins were found using BLASTP (Table 1). Among them, 47 had been previously annotated as WRKY proteins in the CocoaGenDB; the other 14 proteins identified by BLASTP had been previously annotated as uncharacterized, predicted or hypothetical proteins (data not shown). Among the 61 non-redundant sequences, 60 contained at least one complete heptapeptide WRKY motif while one sequence (Tc02_g012180) did not present such a conserved stretch (Table 1). The TcWRKY proteins were classified into three groups according to the presence of WRKY motif and the zinc-finger motif type. Group I contained two WRKY motifs (one in the N-terminal region of the sequence, the other in the C-terminal region) and two C₂H₂ zinc-finger motifs; this group contained 10 TcWRKY proteins (Table 1; Fig 1A). Group II contained only one WRKY motif and a C₂H₂ zinc-finger motif (40 TcWRKY proteins). Group III contained only one WRKY motif and a C₂HC zinc-finger motif (8 TcWRKY proteins) (Table 1, Fig 1A). It was not possible to classify the three other putative WRKY proteins (Tc02_g017240, Tc02_g001170 and Tc02_g012180) because of the presence of a highly altered WRKY motif and/or zinc-finger motif, or because of the absence of a WRKY motif (Table 1; Fig 1A). The zinc finger motifs C₂H₂ from the group I were CX₄CX₂₂HX₁H (N-terminal) and CX₄CX₂₃HX₁H (C-terminal) for all sequences (Fig 2). Group II was divided into five subgroups (IIa to e) according to C₂H₂ zinc-finger structure. Subgroups IIa, IIb, IIc, IId and IIe were found to contain 3, 8, 17, 6 and 6 genes, respectively (Table 1) and the members of subgroups IIa, IId, IIe and 6 members of subgroup IIb showed the CX₅CX₂₃HX₁H zinc-finger motif. The other two members of subgroup IIb (Tc02_g033950 and Tc04_g007790) showed the CX₅CX₂₅HX₁H and CX₅CX₃₁HX₁H zinc-finger structures, respectively (Fig 2). All the members of group IIc showed the CX₄CX₂₃HX₁H zinc-finger structure (Fig 2). In the case of group III, the zinc-finger motif was CX₇CX₂₃HX₁C (Fig 2). The WRKY domain (WD) was highly conserved in 52 proteins, but some of them presented variations (Table 1; Fig 2). The proteins of the group IIc showed a WRKY motif with only one amino acid modification; the protein Tc02_g001230 showed a WRKY motif with two amino acid modifications (WRKHGQT) while Tc00_g017240 contained a WRKY motif with three modifications (WRCIGIK) in addition to the presence of an incomplete zinc-finger motif (Table 1, Fig 2). These seven proteins belong to subgroup IIc, III, or were non-classified (Table 1). The sequence Tc02_g012180, which did not contain any WRKY motif, was removed from all subsequent analyses, while the sequences Tc02_g001170 and Tc02_g012180, which showed modified WRKY motif and modified zinc-finger motif, were excluded from phylogenetic analysis.

Distribution of WRKY genes in the cacao genome

The 61 putative WRKY sequences were distributed on the 10 cacao chromosomes (Fig 1A and 1B). A higher abundance of WRKY genes was observed on chromosome 1: 14 genes belonging to groups I (1 gene), II (12) and III (1) (Fig 1A and 1B). In contrast, chromosomes 8 and 10 contained only two WRKY genes each (from groups II and III, respectively; Fig 1A). The other WRKY genes were distributed as follows: 8, 7, 5, 5, 5, 3 and 6 on chromosomes 2, 3, 4, 5, 6, 7 and 9, respectively, and belonged mainly to groups I and II (Fig 1A). For four of the genes (Tc00_g047270, Tc00_g076580, Tc00_017270, Tc00_g017240), the location was uncertain, so the genes were distributed on "chromosome 00", corresponding to non-anchored sequences of the genome (Table 1; Fig 1A). Because tandem and segmental duplication play an important role in the expansion of multigene families, we analyzed the syntenic regions and structural

changes of all 10 cacao chromosomes (Fig 3). Twenty-four *WRKY* genes were identified in segmental duplication events in the cacao genome (Fig 3). Tc01_g035330 participated in two duplication events with Tc03_g019750 and Tc03_g028030, while Tc01_g034680 also participated in two duplication events with Tc03_g017550 and Tc03_g028700. *TcWRKY* genes

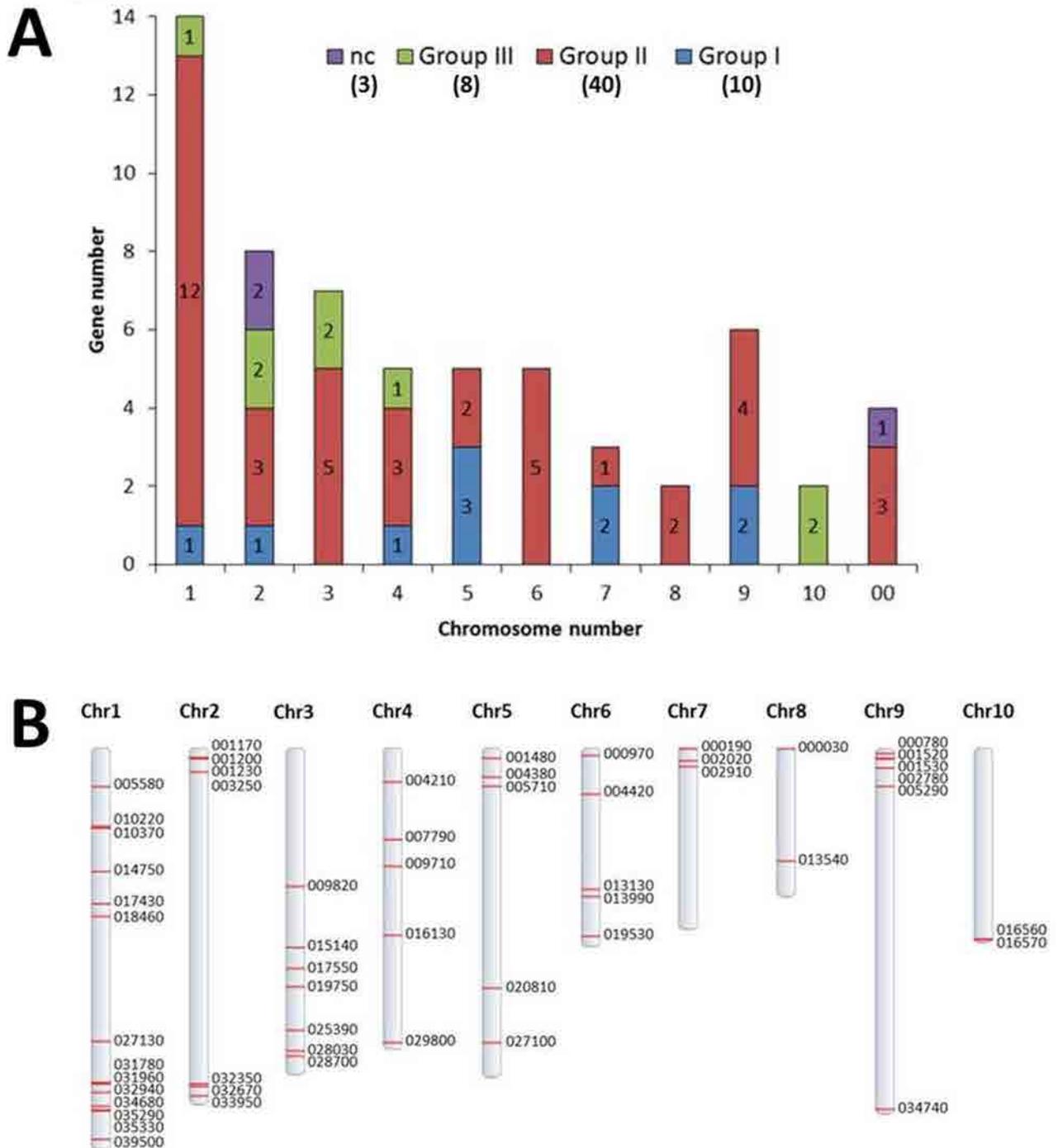


Fig 1. Classification of *TcWRKY* genes and distribution on cacao chromosomes. **A.** Distribution of *TcWRKY* gene groups among cacao chromosomes. Chromosome 00 corresponds to non-anchored genome regions. nc: not classified. The total number of genes in each group was indicated under parenthesis. **B.** Physical distribution of the *TcWRKY* genes on cacao chromosomes. In CocoaGenDB, the sequence names are preceded by ^aTc_g° (e.g., Tc_g005580).

<https://doi.org/10.1371/journal.pone.0187346.g001>

introduced to get the best alignment are indicated by (-). Within each group, (*) represents identical amino acids, (.) and (:) represent conserved substitutions and semi-conserved substitutions, respectively.

<https://doi.org/10.1371/journal.pone.0187346.g002>

were located within syntenic blocks of all chromosomes except chromosome 10. Most of the duplications were located in chromosomes 1 and 3 (Fig 3).

Phylogenetic analysis of TcWRKY proteins

A phylogenetic tree of WRKY amino acid domains from cacao and *Arabidopsis* (S1 Table) was constructed to investigate the relationship between these two species (Fig 4). The WRKY protein domains were grouped according to the general WRKY classification (group I, IIa-e, III). The group I was subdivided according to the N-terminal and C-terminal WD (I* and I**, respectively). The least consistent group was the IIc that was split in two branches, one more related to the group III, the other to the group I**. Groups IIa and IIb presented a close phylogenetic relation, as well as IIc and IIe (Fig 4). The phylogenetic tree constructed with cacao and *Arabidopsis* WRKY domains also allowed inferences to be made about the possible function of cacao sequences based on *Arabidopsis* sequence function knowledge. Clades or sequences associated with plant responses to pathogen or defense inducers (such as abscisic acid or salicylic acid) were identified and cacao sequences with possible defense-related responses were selected for expression analysis (Fig 4). Similar results were obtained by analysis of orthology between cacao and *A. thaliana*: Tc01_g018460 (XP_007049282.1) was orthologue to AT2G04880.2 (WRKY3); Tc04_g016130 (XP_007033512.1) to AT3G56400.1 (WRKY70); Tc10_g016570 (XP_007011367.1) to AT3G56400.1 (WRKY70); Tc09_g001530 (XP_007011727.1) to AT4G31800.2 (WRKY18); Tc06_g004420 (XP_007023430.1) to AT1G80840.1 (WRKY40); Tc06_g013130 (XP_007025165.1) to AT4G18170.1 (WRKY28); and Tc01_g014750 (XP_007048873.1) to AT1G29860.1 (WRKY28).

To investigate the relationship between WRKY family members in cacao more thoroughly, we analyzed the motif pattern of the TcWRKY sequences (Fig 5A). The different motifs were identified based on the biochemical properties of their amino acids as well as their specific location in the protein sequence [52]. The conserved amino acids, the position of each residue in the WRKY sequence, as well as the residue that varied according to the protein sequence are presented in Fig 5A. Twenty motifs were found and 3 of them (motifs 1 to 3; Fig 5B) constituted the WD. Motifs 1 and 2, corresponding to the C-terminal WRKY and the C₂H₂ motifs, were present in 58 TcWRKY members. Motif 3 corresponded to the N-terminal WRKY motif (10 members). Motif 4 is an intermediary amino acid region between motifs 1 and 2, forming the complete WD with approximately 60 amino acids. This motif is present in 58 TcWRKY proteins (Fig 5A and 5B). Six different combinations containing the WRKY motifs were found in the proteins identified in the cacao genome (Fig 5C). Twenty-nine presented the WRKY motif in the center of the sequence and 18 in the C-terminal region. Of the 10 members of group I (with duplicated domains), 9 members presented the WRKY motif in the central region of the protein while one presented this domain in the C-terminal region. Three other members (Tc02_g001230, Tc02_g001200 e Tc02_g001170) presented the WRKY motif domain in the N-terminal region of the protein and also contained the 2 LRR motifs and one NB-ARC domain. One sequence, Tc05_g005710, contained 2 LRR motifs (Fig 5C).

Expression patterns of seven selected TcWRKY genes in resistant and susceptible *Theobroma cacao* genotypes

According to the phylogenetic analysis (Fig 4), the orthology between cacao and *Arabidopsis* and the putative function of WRKY genes in *Arabidopsis* (related to response to biotic stress), 7

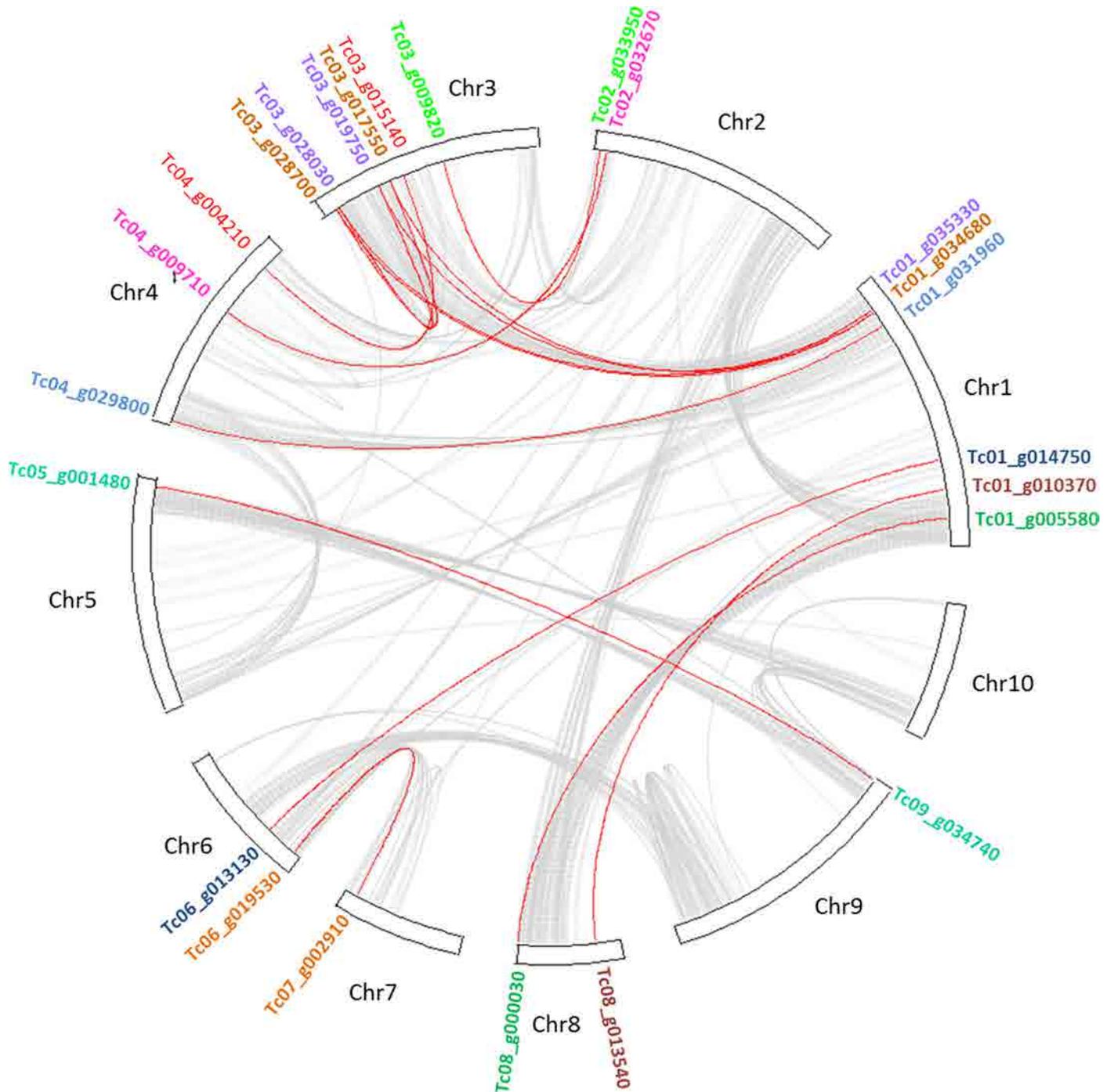


Fig 3. Schematic representation of interchromosomal relationships of *TcWRKY* genes. Gray lines indicate all syntenic blocks in the cacao genome, whereas the red lines suggest duplicated *WRKY* gene pairs. The corresponding *WRKY* gene names are indicated, duplicated genes are marked with the same color. The chromosome number is indicated at the top of each chromosome.

<https://doi.org/10.1371/journal.pone.0187346.g003>

TcWRKY (*Tc04_g016130*, *Tc10_g016570*, *Tc09_g001530*, *Tc06_t004420*, *Tc06_t013130*, *Tc01_t014750* and *Tc01_t018460*) genes were selected for expression analysis by RT-qPCR. From the 7 *TcWRKY* genes analyzed, one belonged to group I (*Tc01_t018460*), two to

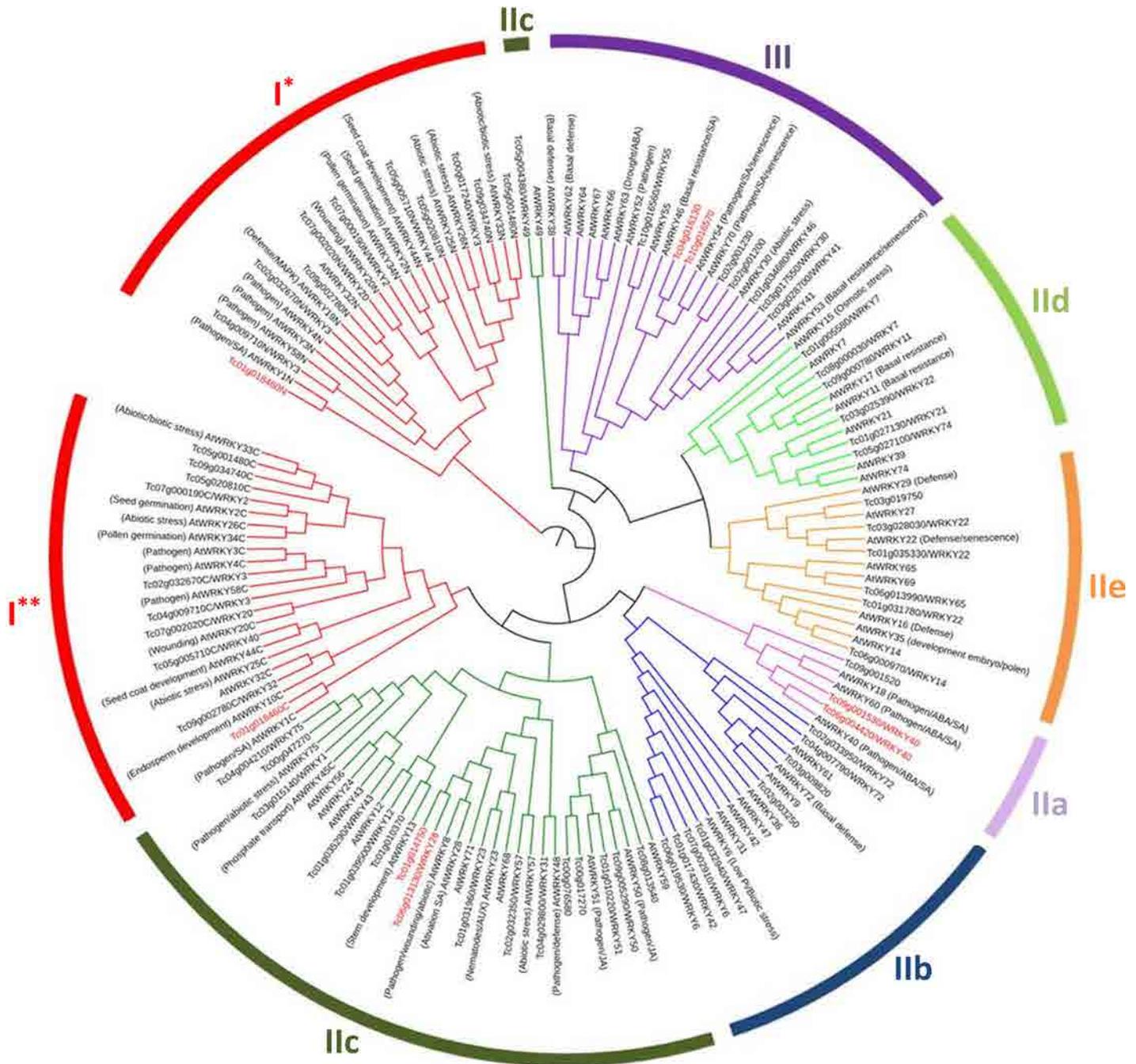


Fig 4. Phylogenetic tree of WRKY protein domains from cacao and Arabidopsis. TcWRKY protein domains (S1 Table) were grouped into three groups and their subgroups as follows: group I in red, subgroup IIa in light purple, subgroup IIb in blue, subgroup IIc in dark green, subgroup IIId in light green, subgroup IIe in orange, group III in dark purple. (*) and (***) indicate the N-terminal and C-terminal WD from group I genes. Cacao WRKY proteins possibly involved in plant defense response and selected for gene expression analysis are indicated in red. These proteins are: Tc04_g016130, Tc10_g016570, Tc09_g001530, Tc06_t004420, Tc06_t013130, Tc01_t014750 and Tc01_t018460 (appearing in the I* and I** groups).

<https://doi.org/10.1371/journal.pone.0187346.g004>

subgroup IIa (Tc06_t004420, Tc09_g001530), two to subgroup IIc (Tc01_t014750, Tc06_t013130) and two to group III (Tc04_g016130, Tc10_g016570). The expression of the TcWRKY genes was analyzed in two cacao genotypes, TSH1188 (resistant to witches' broom

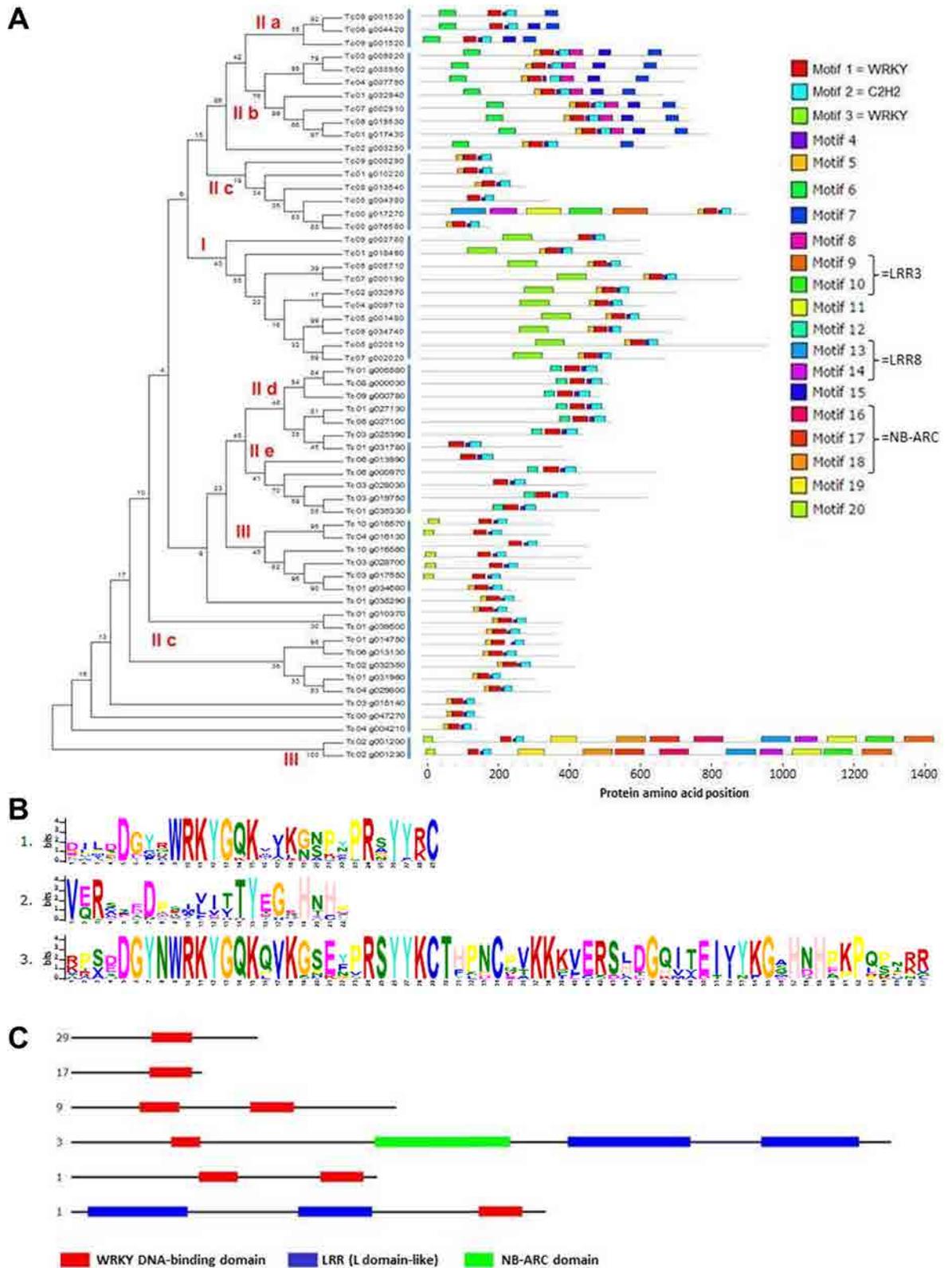


Fig 5. Phylogenetic tree and motif composition of TcWRKY proteins. A. Phylogenetic tree (left side of the figure) and motif composition of the TcWRKY proteins (right side), obtained using the MEGA v.5.1 and the MEME programs, respectively. The

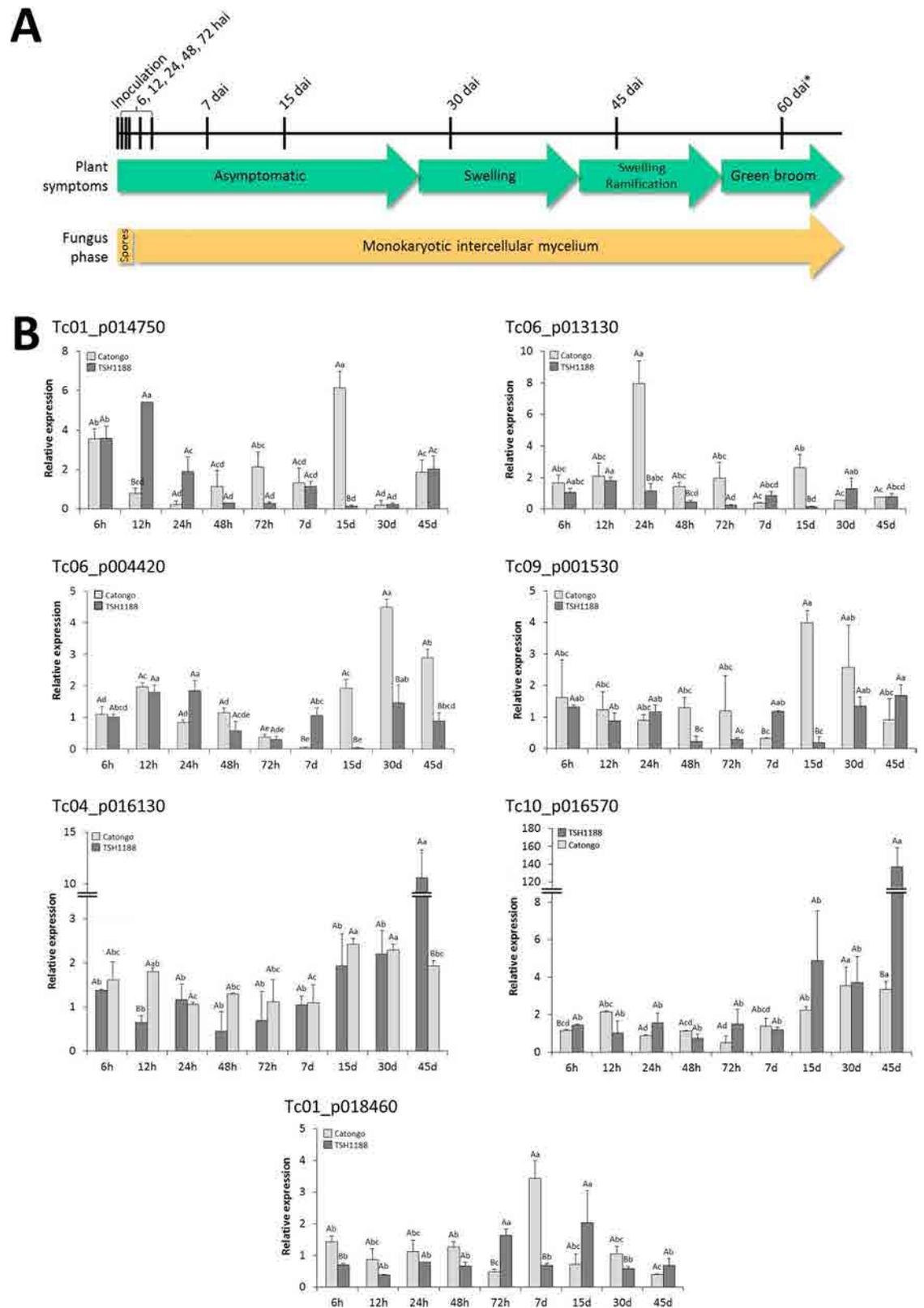
phylogenetic tree contains 58 TcWRKY proteins (excluding Tc00_g017240, Tc02_g001170 and Tc02_g012180). WRKY groups are indicated in red. **B.** Motif detail of the three first most probable motifs (WRKY C-terminal, C₂H₂ and WRKY N-terminal motifs) of the TcWRKY proteins, obtained by the MEME program. **C.** Organization of the 60 TcWRKY proteins (excluding Tc02_g012180) in 6 categories by domain combination using the SuperFamily database. The numbers indicated on the left represent the number of cacao WRKY proteins in each category.

<https://doi.org/10.1371/journal.pone.0187346.g005>

disease) and Catongo (susceptible) infected or not (control) with *M. perniciosa* (Fig 6A and 6B). For both genotypes and for all the harvesting points, the PCR amplification occurred at the same melting temperature, showing that only the target gene was amplified (data not shown). RT-qPCR analysis showed differential expression between genotypes and/or between time intervals for all the analyzed genes. In Catongo, the *TcWRKY* genes Tc06_p004420, Tc09_p001530, Tc04_p016130 and Tc10_p016570 showed higher transcript abundance in the final stages of the infection (15 to 45 dai). The gene Tc01_p014750 showed higher transcript abundance 15 dai but also at the beginning of the infection (6 hai; about 6 times more expressed). The gene Tc06_p013130 showed higher transcript abundance (about 7 times more than the control) 24 hai while the gene Tc01_p018460 was more expressed 7 dai (Fig 6B). In TSH1188, the transcript abundance of the genes Tc06_p013130, Tc06_p004420, Tc09_p001530 and Tc01_p018460 was low (about 2 times more expressed than control). The gene Tc01_p014750 was mainly expressed in the initial infection points (6 to 24 hai; about 5 times more expressed). Interestingly, the Tc04_p016130 and Tc10_p016570 *TcWRKY* genes showed very high transcript abundance 45 dai: Tc04_p016130 showed an increase of about 12 times while Tc10_p016570 was expressed 120 times more than the control (Fig 6B). Some significant differences were also observed between genotypes: 12 hai and 15 dai for the gene Tc01_p014750; 24 hai, 48 hai and 15 dai for the gene Tc06_p013130; 7, 15, 30 and 45 dai for the gene Tc06_p004420; 48 hai, 7 and 15 dai for the gene Tc09_p001530; 12 hai and 45 dai for the gene Tc04_p016130; 6 hai and 45 dai for the gene Tc10_p016570; and 6 and 72 hai, 7 and 30 dai for the gene Tc01_p018460 (Fig 6B).

Discussion

WRKY proteins constitute one of the most important transcription factor families in plants due to their participation in diverse biological processes, including responses to biotic and abiotic stresses [7]. A better understanding of this family, including member characterization, phylogenetic analysis and expression analysis, can help to define new disease management strategies, as in the case of the cacao-*M. perniciosa* interaction. In this study, based on sequence comparison and molecular phylogeny, 58 cacao proteins with complete WD were found (Table 1). These proteins belonged to three main WRKY groups and their subgroups, which were distributed throughout the genome (Fig 1). Seven members presented variations in the WD (in the heptapeptide or in the C₂H₂ zinc finger motifs), suggesting a higher divergence, possibly due to recent mutations, of these genes in comparison to the rest of the TcWRKY family. Groups IIc and III (5 and 1 gene, respectively) contained 70% of the amino acid variations observed, suggesting that these two groups were more subjected to selective pressure and variability through time. Analysis in other species such as cotton or tomato, also showed that these two groups were the most divergent in the evolutionary history of the WRKY family [53, 54]. The phylogenetic analysis revealed that subgroups IIa and IIb are sister groups and share a common ancestor, as well as subgroups IIc and IIe (Fig 4). Various studies have demonstrated that the expansion of WRKY TF family members is mainly due to gene duplication events, as shown in rice [55], *Arabidopsis* [56], cotton [53], *Populus* [57] and barley [58]. Here, we observed that 40% of the *TcWRKY* sequences presented one or more duplication events (Fig 3) and that these events were associated mainly with the conservation of the TcWRKY motif



course in Catongo genotype. The harvesting times of inoculated and control plants are indicated on the top of the figure, excepted (*) that was used only of plant symptoms observation. **B.** RT-qPCR of *TcWRKY* genes. The control used as calibrator in the expression value calculation corresponds to the control plants (mock-inoculated with water) collected at each harvesting time and used as calibrators of the corresponding inoculated sample (see also [Methods](#) section). The results are the arithmetical mean of the repetitions \pm standard error. Different lower case letters indicate significant statistical difference between harvesting times for each genotype by the Duncan test ($P \leq 0.05$), while upper case letters correspond to significant statistical difference between genotypes for each harvesting time by *t*-test ($P \leq 0.05$). d: days after inoculation; h: hours after inoculation.

<https://doi.org/10.1371/journal.pone.0187346.g006>

patterns (Figs 4 and 5). Generally, the duplicated genes were also present together in the same clades of the motif phylogeny as observed for Tc04_g029800/Tc01_g031960, Tc04_g009710/Tc02_g032670, Tc03_g028700/Tc03_g017550/Tc01_g034680, Tc03_g009820/Tc02_g033950, Tc03_g028030/Tc03_g019750/Tc01_g035330, Tc01_g014750/Tc06_g013130, Tc07_g002910/Tc06_g019530, Tc09_g034740/Tc05_g001480, Tc04_g004210/Tc03_g015140 and Tc08_g000030/Tc01_g005580 (Figs 3, 4 and 5). However, in the case of the duplication of Tc08_g013540/Tc01_g010370, the two sequences were located in different phylogenetic clades, suggesting an evolution of the gene and motif structure (Figs 3 and 4). Phylogenetic analysis in macromolecules, by forming non-random clusters, also suggests that these molecules may share the same biological functions, may be present in the same cell compartment or be expressed/produced at the same moment during a biological process. Besides the heptapeptide WRKYGQK and the C₂H₂/C₂HC zinc-finger motifs \pm known to be involved in DNA-binding \pm some *TcWRKY* proteins (Tc02_g001230, Tc02_g001200, Tc02_g001170 and Tc00_g017270) showed conserved motifs, such as NB-ARC and/or LRR (Fig 5C), known to be involved in pathogen recognition, plant resistance and activation of plant immunity [59, 60].

Expression and functional analysis of WRKY TFs could help in discriminating the role and function of these proteins at the tissue and organism levels. Here, we evaluated by RT-qPCR the expression of seven *TcWRKY* genes in resistant and susceptible cacao plants inoculated or not with *M. perniciosa*. The choice of the genes was based on previous indications in the genome databank (CocoaGenDB), phylogenetic analysis and putative function of the orthologues in *Arabidopsis*, showing that: i) Tc01_p014750 and Tc06_p013130 (indicated as *TcWRKY28*) were co-orthologous to *AtWRKY8* and *AtWRKY28*; ii) Tc09_p001530 and Tc06_p004420 (both indicated as *TcWRKY40*) were both orthologous to *AtWRKY40*; iii) Tc04_p016130 and Tc10_p016570 were co-orthologous to *AtWRKY54* and *AtWRKY70*; and iv) Tc01_p018460 was orthologous to *AtWRKY1* (Fig 4). Interestingly, the Tc01_p014750 and Tc06_p013130 genes came from an event of duplication (Fig 3) but showed different expression patterns (Fig 6B), suggesting that sequence evolution may result in different roles and/or functions in relation to pathogen response. These two sequences were co-orthologous to *AtWRKY8* and *AtWRKY28*, which are induced by ABA, wounding, oxalic acid (OA) and/or hydrogen peroxide (H₂O₂) [61, 62]. In cacao, it has been reported that the amount of calcium oxalate crystal (COC) and H₂O₂ levels in the TSH1188 (resistant) vs. Catongo (susceptible) varieties present distinct temporal and genotype dependent patterns [40, 41]: susceptible variety accumulated more COC than the resistant one, and the COC dissolution \pm resulting in OA and H₂O₂ formation \pm occurred in the early infection stages in the resistant variety and in the final stage of the disease in the susceptible one. Interestingly, the *Tc01_p014750* gene, orthologous to *AtWRKY28*, showed higher expression in the early infection stages in TSH1188 and 15 dai in the Catongo variety (Fig 6B), stages during which the H₂O₂ is considered the highest [40, 41]. The *Tc06_p013130* gene (previously annotated as *TcWRKY28*) showed an expression pattern different from Tc01_p014750, with high abundance 24 hai in the susceptible variety and a constant and low abundance in the resistant one (Fig 6). Such divergent behavior was

previously observed for the rice *WRKY28* gene. Delteil et al. [63] reported that the knock-out of *OsWRKY28* by T-DNA insertion leads to a two-fold increase in resistance to a compatible rice blast fungus, and this phenotype is accompanied with increased expression of several defense-related genes. Likewise, other authors showed that the overexpression of *OsWRKY28* resulted in enhanced susceptibility to the rice blast fungus *Magnaporthe oryzae* and decreased accumulation of PR-5 [64]. According to the authors, these phenotypes observed in overexpression or genetic defects in *OsWRKY28* are consistent with their presumed role as negative regulators of basal defense responses to compatible rice blast fungus strains [64]. The same role was also suggested for *WRKY8* in *Arabidopsis*. This gene could be a negative or positive regulator of the basal resistance of the plant when infected by *Pseudomonas syringae* or *Botrytis cinerea*, respectively [61]. In cacao, it can be suggested that Tc01_p014750 acted as a positive regulator of plant resistance to *M. perniciosa* through activation by OA and/or reactive oxygen species (ROS); the involvement of OA and ROS in cacao resistance to *M. perniciosa* has been previously observed in studies using the same or similar plant genotypes and culture conditions [40, 41, 65]. On the other hand, Tc06_p013130 may have acted as negative regulator of the basal resistance of cacao.

The Tc09_p001530 and Tc06_p004420 genes (both indicated as *TcWRKY40*) showed high transcript abundance in the susceptible cacao variety (final stages; Fig 6B). In *Arabidopsis*, studies have shown that the WRKY18, WRKY40 and WRKY60 TFs are induced by pathogens and interact physically and functionally together forming homo and heterocomplexes [66]. The constitutive overexpression of these genes in *Arabidopsis* increased its susceptibility to *B. cinerea* [66]. Moreover, the superexpression of WRKY40 in transgenic *Populus trichocarpa* plants conferred high susceptibility to the hemibiotrophic fungus *Dothiorella gregaria* Sacc., indicating that PtrWRKY40 plays a negative role in resistance to this hemibiotrophic fungus in poplar [67]. In cacao, the Tc09_p001530 and Tc06_p004420 genes (*TcWRKY40*) may have a similar function to that observed in poplar: the gene expression in the Catongo variety may be associated with the plant susceptibility to *M. perniciosa*. In the phylogenetic analysis, the sequences Tc04_p016130 and Tc10_p016570 were grouped both with AtWRKY54 and AtWKY70 (Fig 4). Tc04_p016130 and Tc10_p016570 showed similar expression patterns, mainly with a very high expression in the last time point in TSH1188 (about 10 and 140 times more at 45 dai, respectively). Several works have reported the cooperation of the *AtWRKY54* and *AtWKY70* genes in response to biotic and abiotic stresses [68, 69]. These TFs are positive regulators of plant defense, and cooperate as negative regulators of salicylic acid (SA) biosynthesis and senescence [69], but are not responsive to signals such as ROS [68]. The WRKY70 TF was identified as an integrator in cross-talk between SA and jasmonic acid (JA), two hormones with a well-defined function in plant defense response regulation [70]. Generally, SA is associated with defense response against biotrophic pathogens, whereas JA has a function in defense responses against herbivore and necrotrophic pathogens [70]. In cacao, previous works have shown an increase of jasmonate biosynthesis genes in the last time stages in TSH1188-*M. perniciosa* interaction (from 30 to 60 dai), as well as an increase of ROS detoxification genes [39]. The gene Tc01_p018460 showed high phylogenetic proximity with AtWRKY1 but also with other genes involved in pathogen responses induced by SA, such as AtWRKY3 and ATWRKY58 (Fig 4); this proximity may be related to similar gene function. In the susceptible genotype, the *Tc01_p018460* expression was higher mainly 7 dai (Fig 6B).

Conclusion

Here, we identified 61 WRKY proteins from *T. cacao*, distributed on all the chromosomes, in some cases coming from different duplication events. To our knowledge, this is the first report

of the entire WRKY TF family in cacao and of expression analysis in relation to *M. perniciosa* infection. The TcWRKY family showed a phylogenetic composition similar to that of *Arabidopsis* and some couple of sequences showed similar expression patterns and possibly functions (e.g., Tc01_p014750/Tc06_p013130/AtWRKY28; Tc09_p001530/Tc06_p004420/AtWRKY40; Tc04_p016130/AtWRKY54; Tc10_p016570/AtWRKY70; Tc01_p018460/AtWRKY1). Mainly, the Tc04_p016130 and Tc10_p016570 sequences presented a special interest due to their high and differential expression level between resistant and susceptible plants infected by *M. perniciosa*. In general, our results can help to select appropriate candidate genes for further characterization and/or confirmation studies in relation to pathogen resistance in cacao or in other *Theobroma* species, as well as for support of future breeding efforts.

Supporting information

S1 Fig. General scheme of the *in silico* pipeline used for identification and confirmation of the TcWRKY sequences.

(DOCX)

S2 Fig. Disease symptoms observed in the TSH1188 (resistant) and Catongo (susceptible) infected and non-infected plants. White arrow: swelling of the stem; black arrow: ramification (green broom).

(DOCX)

S1 Table. List of the 153 *T. cacao* and *A. thaliana* WRKY domains used for phylogeny.

(DOCX)

S2 Table. Primers used in this study.

(DOCX)

S3 Table. Characteristics of WRKY amplicons.

(DOCX)

Acknowledgments

We thank Dr. Aurizângela Oliveira de Sousa and Dr. Sara Pereira Menezes (UESC) for helpful comments during RT-qPCR design and Dr. Claudia Fortes Ferreira (Embrapa CNPMF, Brazil) for the English language revision.

Author Contributions

Conceptualization: Dayanne Silva Monteiro de Almeida, Daniel Oliveira Jordão do Amaral, Michel Vincentz, Fabienne Micheli.

Formal analysis: Dayanne Silva Monteiro de Almeida, Luiz-Eduardo Del-Bem, Raner José Santana Silva, Michel Vincentz.

Funding acquisition: Fabienne Micheli.

Investigation: Dayanne Silva Monteiro de Almeida, Daniel Oliveira Jordão do Amaral, Emily Bronze dos Santos, Raner José Santana Silva, Karina Peres Gramacho.

Methodology: Daniel Oliveira Jordão do Amaral, Luiz-Eduardo Del-Bem, Karina Peres Gramacho.

Project administration: Fabienne Micheli.

Resources: Karina Peres Gramacho, Fabienne Micheli.

Supervision: Michel Vincentz, Fabienne Micheli.

Validation: Fabienne Micheli.

Visualization: Dayanne Silva Monteiro de Almeida, Fabienne Micheli.

Writing ± original draft: Dayanne Silva Monteiro de Almeida, Fabienne Micheli.

Writing ± review & editing: Dayanne Silva Monteiro de Almeida, Karina Peres Gramacho, Fabienne Micheli.

References

1. Kole C, Muthamilarasan M, Henry R, Edwards D, Sharma R, Abberton M, et al. Application of genomics-assisted breeding for generation of climate resilient crops: Progress and prospects. *Front Plant Sci.* 2015; 6:1±16. <https://doi.org/10.3389/fpls.2015.00001>
2. Mickelbart MV, Hasegawa PM, Bailey-Serres J. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nat Rev Genet.* 2015; 16(4):237±251. <https://doi.org/10.1038/nrg3901> PMID: 25752530
3. Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Prasad M. Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in Foxtail Millet (*Setaria italica* L.). *PLoS ONE.* 2014; 9(11):e113092. <https://doi.org/10.1371/journal.pone.0113092> PMID: 25409524
4. Burley SK, Kamada K. Transcription factor complexes. *Curr Opin Struct Biol.* 2002; 12(2):225±230. [https://doi.org/10.1016/S0959-440X\(02\)00314-7](https://doi.org/10.1016/S0959-440X(02)00314-7) PMID: 11959501
5. Ishiguro S, Nakamura K. 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 beta-amylase from sweet potato. *Mol Gen Genet.* 1994; 244(6):563±571. PMID: 7969025
6. Eulgem T, Rushton P, Robatzek S, Somssich I. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 2000; 5:199±206. [https://doi.org/10.1016/s1360-1385\(00\)01600-9](https://doi.org/10.1016/s1360-1385(00)01600-9) PMID: 10785665
7. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends Plant Sci.* 2010; 15(5):247±258. <https://doi.org/10.1016/j.tplants.2010.02.006> PMID: 20304701
8. Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol.* 2005; 5(1):1±12. <https://doi.org/10.1186/1471-2148-5-1> PMID: 15629062
9. Zhang Y, Feng Jc. Identification and characterization of the grape WRKY Family. *Biomed Res Int.* 2014; 2014:1±14. <https://doi.org/10.1155/2014/787680> PMID: 24883326
10. Cai H, Yang S, Yan Y, Xiao Z, Cheng J, Wu J, et al. CaWRKY6 transcriptionally activates CaWRKY40, regulates *Ralstonia solanacearum* resistance, and confers high-temperature and high-humidity tolerance in pepper. *J Exp Bot.* 2015; 66(11):3163±3174. <https://doi.org/10.1093/jxb/erv125> PMID: 25873659
11. Lagac e M, Matton DP. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta.* 2004; 219(1):185±189. <https://doi.org/10.1007/s00425-004-1253-2> PMID: 15045588
12. Johnson CS, Kolevski B, Smyth DR. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *Plant Cell.* 2002; 14:1359±1375. <https://doi.org/10.1105/tpc.001404> PMID: 12084832
13. Miao Y, Laun T, Zimmermann P, Zentgraf U. Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. *Plant Mol Biol.* 2004; 55:853±867. <https://doi.org/10.1007/s11103-004-2142-6> PMID: 15604721
14. Guan Y, Meng X, Khanna R, LaMontagne E, Liu Y, Zhang S. Phosphorylation of a WRKY transcription factor by MAPKs is required for pollen development and function in *Arabidopsis*. *PLoS Genet.* 2014; 10(5):e1004384. <https://doi.org/10.1371/journal.pgen.1004384> PMID: 24830428
15. Rinerson CI, Scully ED, Palmer NA, Donze-Reiner T, Rabara RC, Tripathi P, et al. The WRKY transcription factor family and senescence in switchgrass. *BMC Genomics.* 2015; 16:1±17. <https://doi.org/10.1186/1471-2164-16-1>
16. Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L. The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol.* 2011; 157(4):2081±2093. <https://doi.org/10.1104/pp.111.181834> PMID: 21988879
17. Zhang ZL. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiol.* 2004; 134(4):1500±1513. <https://doi.org/10.1104/pp.103.034967> PMID: 15047897

18. Pandey SP, Roccaro M, Schenck M, Logemann E, Somssich IE. Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *Plant J.* 2010; 64(6):912±923. <https://doi.org/10.1111/j.1365-313X.2010.04387.x> PMID: 21143673
19. Mukhtar MS, Deslandes L, Auriac M-C, Marco Y, Somssich IE. The Arabidopsis transcription factor WRKY27 influences wilt disease symptom development caused by *Ralstonia solanacearum*. *Plant J.* 2008; 56(6):935±947. <https://doi.org/10.1111/j.1365-313X.2008.03651.x> PMID: 18702671
20. Bhattarai KK, Atamian HS, Kaloshian I, Eulgem T. WRKY72-type transcription factors contribute to basal immunity in tomato and Arabidopsis as well as gene-for-gene resistance mediated by the tomato R gene *Mi-1*. *Plant J.* 2010; 63(2):229±240. <https://doi.org/10.1111/j.1365-313X.2010.04232.x> PMID: 20409007
21. Lopes MA, Hora BT Jr., Dias CV, Santos GC, Gramacho KP, Cascardo JC, et al. Expression analysis of transcription factors from the interaction between cacao and *Moniliophthora perniciosa* (Tricholomataceae). *Genet Mol Res.* 2010; 9(3):1279±1297. <https://doi.org/10.4238/vol9-3gmr825> PMID: 20623454
22. Borrone JW, Meerow AW, Kuhn DN, Whitlock BA, Schnell RJ. The potential of the WRKY gene family for phylogenetic reconstruction: An example from the Malvaceae. *Mol Phylogenet Evol.* 2007; 44(3):1141±1154. <https://doi.org/10.1016/j.ympev.2007.06.012> PMID: 17681475
23. Borrone JW, Kuhn DN, Schnell RJ. Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theor Appl Genet.* 2004; 109(3):495±507. <https://doi.org/10.1007/s00122-004-1662-4> PMID: 15148571
24. Aime MC, Phillips-Mora W. The causal agents of witches' broom and frosty pod rot of cacao (*Theobroma cacao*) form a new lineage of Marasmiaceae. *Mycologia.* 2005; 97(5):1012±1022. <https://doi.org/10.3852/mycologia.97.5.1012> PMID: 16596953
25. Purdy L, Schmidt R. Status of cacao witches' broom: biology, epidemiology, and management. *Annu Rev Phytopathol.* 1996; 34(1):573±594. <https://doi.org/10.1146/annurev.phyto.34.1.573> PMID: 15012557
26. Griffith GW, Nicholson J, Nenninger A, Birch RN, Hedger JN. Witches' brooms and frosty pods: Two major pathogens of cacao. *New Zealand Journal of Botany.* 2003; 41(3):423±435. <https://doi.org/10.1080/0028825x.2003.9512860>
27. Sayid R. Chocolate Could Run Out In 2020 Due To Worldwide Shortage of Cocoa. *The Daily Mirror online.* 2013 31/12/2013; Sect. World News.
28. Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, et al. The genome of *Theobroma cacao*. *Nat Genet.* 2011; 43:101±108. <https://doi.org/10.1038/ng.736> PMID: 21186351
29. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012; 40(7):e49±e49. <https://doi.org/10.1093/nar/gkr1293> PMID: 22217600
30. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 2009; 37(Web Server issue):W202±W208. <https://doi.org/10.1093/nar/gkp335> PMID: 19458158
31. Gough J, Karplus K, Hughey R, Chothia C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *Journal of Molecular Biology.* 2001; 313(4):903±919. <https://doi.org/10.1006/jmbi.2001.5080> <https://doi.org/10.1006/jmbi.2001.5080> PMID: 11697912
32. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res.* 2004; 32:1792±1797. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
33. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011; 28(10):2731±2739. <https://doi.org/10.1093/molbev/msr121> PMID: 21546353
34. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987; 4(4):406±425. PMID: 3447015
35. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 1998; 14(9):817±818. <https://doi.org/10.1093/bioinformatics/14.9.817> PMID: 9918953
36. Lopes UV, Monteiro WR, Pires JL, Clement D, Yamada MM, Gramacho KP. Cacao breeding in Bahia, Brazil: Dst strategies and results. *Crop Breed Appl Biotechnol.* 2011; 1:73±81.
37. Pires JL, De Melo GP, Yamada MM, Gramacho KP. Association among sources of resistance to witches' broom disease for the increment of the level and durability of the character. *Agrotropica* 2012; 24:27±30.

38. Gesteira AS, Micheli F, Carels N, Da Silva A, Gramacho K, Schuster I, et al. Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa*. *Ann Bot.* 2007; 100(1):129±140. <https://doi.org/10.1093/aob/mcm092> PMID: 17557832
39. da Hora Junior BT, Poloni JdF, Lopes MA, Dias CV, Gramacho KP, Schuster I, et al. Transcriptomics and systems biology analysis in identification of specific pathways involved in cacao resistance and susceptibility to witches' broom disease. *Mol Biosyst.* 2012; 8(5):1507±1519. <https://doi.org/10.1039/c2mb05421c> PMID: 22373587
40. de Oliveira Ceita G, Macêdo JNA, Santos TB, Alemanno L, da Silva Gesteira A, Micheli F, et al. Involvement of calcium oxalate degradation during programmed cell death in *Theobroma cacao* tissues triggered by the hemibiotrophic fungus *Moniliophthora perniciosa*. *Plant Sci.* 2007; 173(2):106±117. <https://doi.org/10.1016/j.plantsci.2007.04.006>
41. Dias CV, Mendes JS, dos Santos AC, Pirovani CP, da Silva Gesteira A, Micheli F, et al. Hydrogen peroxide formation in cacao tissues infected by the hemibiotrophic fungus *Moniliophthora perniciosa*. *Plant Physiol Biochem.* 2011; 49(8):917±922. <https://doi.org/10.1016/j.plaphy.2011.05.004> PMID: 21641227
42. Sena K, Alemanno L, Gramacho KP. The infection process of *Moniliophthora perniciosa* in cacao. *Plant Pathology.* 2014:1272±1281. <https://doi.org/10.1111/ppa.12224>
43. Castellani A. Maintenance and cultivation of common pathogenic fungi of man in sterile distilled water. *Further Researches. J Trop Med Hyg.* 1967; 70:181±184.
44. Surujdeo-Maharaj S, Umaharan P, Butler DR, Sreenivasan TN. An optimized screening method for identifying levels of resistance to *Crinipellis perniciosa* in cocoa (*Theobroma cacao*). *Plant Pathol.* 2003; 52(4):464±475. <https://doi.org/10.1046/j.1365-3059.2003.00865.x>
45. Silva S, Luz E, Almeida O, Gramacho K, Bezerra J. Redescoberta da sintomatologia causada por *Crinipellis perniciosa* em cacauero. *Agrotropica* 2002; 1:1±23.
46. Silva S, Matsuoka K. Histologia da interação *Crinipellis perniciosa* em cacaueros suscetíveis e resistentes à vassoura de bruxa. *Fitopatol Bras.* 1999; 24:54±59.
47. Pereira Menezes S, de Andrade Silva E, Matos Lima E, Oliveira de Sousa A, Silva Andrade B, Santos Lima Lemos L, et al. The pathogenesis-related protein PR-4b from *Theobroma cacao* presents RNase activity, Ca²⁺ and Mg²⁺ dependent-DNase activity and antifungal action on *Moniliophthora perniciosa*. *BMC Plant Biol.* 2014; 14(1):1±17. <https://doi.org/10.1186/1471-2229-14-161> PMID: 24920373
48. Ririe KM, Rasmussen RP, Wittwer CT. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal Biochem.* 1997; 245(2):154±160. <https://doi.org/10.1006/abio.1996.9916> PMID: 9056205
49. Pinheiro TT, Litholdo CG Jr., Sereno ML, Leal GA Jr., Albuquerque PSB, Figueira A. Establishing references for gene expression analyses by RT-qPCR in *Theobroma cacao* tissues. *Genet Mol Res.* 2012; 10(4):3291±3305. <https://doi.org/10.4238/2011.November.17.4> PMID: 22095481
50. Zhao S, Fernald RD. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol.* 2005; 12(8):1047±1064. <https://doi.org/10.1089/cmb.2005.12.1047> PMID: 16241897
51. Canteri MG, Althaus RA, das Virgens Filho JS, Gigliotti EA, Godoy CV. SASM-AGRID Sistema para análise e separação de médias em experimentos agrícolas pelos métodos Scott-Knott, Tukey a Duncan. *Revista Brasileira de Agrocomputação.* 2001; 1(2):18±24.
52. Bailey TL, Gribskov M. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics.* 1998; 14(1):48±54. <https://doi.org/10.1093/bioinformatics/14.1.48> PMID: 9520501
53. Cai C, Niu E, Du H, Zhao L, Feng Y, Guo W. Genome-wide analysis of the WRKY transcription factor gene family in *Gossypium raimondii* and the expression of orthologs in cultivated tetraploid cotton. *Crop J.* 2014; 2(2±3):87±101. <https://doi.org/10.1016/j.cj.2014.03.001>
54. Huang S, Gao Y, Liu J, Peng X, Niu X, Fei Z, et al. Genome-wide analysis of WRKY transcription factors in *Solanum lycopersicum*. *Mol Genet Genomics.* 2012; 287(6):495±513. <https://doi.org/10.1007/s00438-012-0696-6> PMID: 22570076
55. Ross CA, Liu Y, Shen QJ. The WRKY gene family in rice (*Oryza sativa*). *J Integr Plant Biol.* 2007; 49(6):827±842. <https://doi.org/10.1111/j.1744-7909.2007.00504.x>
56. Wang Q, Wang M, Zhang X, Hao B, Kaushik SK, Pan Y. WRKY gene family evolution in *Arabidopsis thaliana*. *Genetica.* 2011; 139(8):973±983. <https://doi.org/10.1007/s10709-011-9599-4> PMID: 21805321
57. Jiang Y, Duan Y, Yin J, Ye S, Zhu J, Zhang F, et al. Genome-wide identification and characterization of the Populus WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. *J Exp Bot.* 2014; 65(22):6629±6644. <https://doi.org/10.1093/jxb/eru381> PMID: 25249073

58. Mangelsen E, Kilian J, Berendzen KW, Kolukisaoglu UH, Harter K, Jansson C, et al. Phylogenetic and comparative gene expression analysis of barley (*Hordeum vulgare*) WRKY transcription factor family reveals putatively retained functions between monocots and dicots. *BMC Genomics*. 2008; 9(1):1±17. <https://doi.org/10.1186/1471-2164-9-194> PMID: 18442363
59. van Ooijen G, Mayr G, Kasiem MMA, Albrecht M, Cornelissen BJC, Takken FLW. Structure±function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*. 2008; 59(6):1383±1397. <https://doi.org/10.1093/jxb/ern045> PMID: 18390848
60. Caplan J, Padmanabhan M, Dinesh-Kumar SP. Plant NB-LRR Immune Receptors: From Recognition to Transcriptional Reprogramming. *Cell Host & Microbe*. 3(3):126±135. <https://doi.org/10.1016/j.chom.2008.02.010> PMID: 18329612
61. Chen L, Zhang L, Yu D. Wounding-induced WRKY8 is involved in basal defense in Arabidopsis. *Mol Plant Microbe Interact*. 2010; 23(5):558±565. <https://doi.org/10.1094/MPMI-23-5-0558> PMID: 20367464
62. Chen X, Liu J, Lin G, Wang A, Wang Z, Lu G. Overexpression of AtWRKY28 and AtWRKY75 in Arabidopsis enhances resistance to oxalic acid and *Sclerotinia sclerotiorum*. *Plant Cell Rep*. 2013; 32(10):1589±1599. <https://doi.org/10.1007/s00299-013-1469-3> PMID: 23749099
63. Delleil A, Blein M, Faivre-Rampant O, Guellim A, Estevan J, Hirsch J, et al. Building a mutant resource for the study of disease resistance in rice reveals the pivotal role of several genes involved in defence. *Mol Plant Pathol*. 2012; 13(1):72±82. <https://doi.org/10.1111/j.1364-3703.2011.00731.x> PMID: 21726398
64. Chujo T, Miyamoto K, Shimogawa T, Shimizu T, Otake Y, Yokotani N, et al. OsWRKY28, a PAMP-responsive transrepressor, negatively regulates innate immune responses in rice against rice blast fungus. *Plant Mol Biol*. 2013; 82(1):23±37. <https://doi.org/10.1007/s11103-013-0032-5> PMID: 23462973
65. Fister AS, O'Neil ST, Shi Z, Zhang Y, Tyler BM, Gultinan MJ, et al. Two *Theobroma cacao* genotypes with contrasting pathogen tolerance show aberrant transcriptional and ROS responses after salicylic acid treatment. *Journal of Experimental Botany*. 2015:6245±6258. <https://doi.org/10.1093/jxb/erv334> PMID: 26163705
66. Xu X, Chen C, Fan B, Chen Z. Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell*. 2006; 18(5):1310±1326. <https://doi.org/10.1105/tpc.105.037523> PMID: 16603654
67. Karim A, Jiang Y, Guo L, Ling Z, Ye S, Duan Y, et al. Isolation and characterization of a subgroup IIa WRKY transcription factor PtrWRKY40 from *Populus trichocarpa*. *Tree Physiol*. 2015; 35(10):1129±1139. <https://doi.org/10.1093/treephys/tpv084> PMID: 26423133
68. Besseau S, Li J, Palva ET. WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in *Arabidopsis thaliana*. *J Exp Bot*. 2012; 63(7):2667±2679. <https://doi.org/10.1093/jxb/err450> PMID: 22268143
69. Li J, Besseau S, Tränkle P, Sipari N, Kollist H, Holm L, et al. Defense-related transcription factors WRKY70 and WRKY54 modulate osmotic stress tolerance by regulating stomatal aperture in Arabidopsis. *New Phytol*. 2013; 200(2):457±472. <https://doi.org/10.1111/nph.12378> PMID: 23815736
70. Shim JS, Jung C, Lee S, Min K, Lee Y-W, Choi Y, et al. AtMYB44 regulates WRKY70 expression and modulates antagonistic interaction between salicylic acid and jasmonic acid signaling. *Plant J*. 2013; 73(3):483±495. <https://doi.org/10.1111/tpj.12051> PMID: 23067202