HVA22 from citrus: a small gene family whose some members are involved in plant response to abiotic stress

Monique Drielle Gomes Ferreira[1], Jacqueline Araújo Castro[1,2], Raner José Santana Silva[1], Fabienne Micheli[1,3]

[1] Universidade Estadual de Santa Cruz (UESC), Centro de Biotecnologia e Genética (CBG), Ilhéus, Bahia, Brazil,

[2] Instituto Federal de Educação, Ciência e Tecnologia Baiano (IFBaiano), Governador Mangabeira, Bahia, Brazil,

[3] CIRAD, UMR AGAP, F-34398 Montpellier, France

Abstract

The *HVA22* gene has been isolated for the first time from the aleurone layer of barley (*Hordeum vulgare*). Here, we characterized the HVA22 family from citrus (*C. clementina* and *C. sinensis*). Twelve genes, 6 in each species, were identified as well as duplication events for some of them. The ORF size ranged from 235 to 804 bp and the protein molecular weight from 94 to 267 kDa. All the citrus HVA22 protein presented transmembrane location and conserved TB2/DP1/HVA22 region. Phylogenetic and gene expression analyses suggested that some citrus HVA22 play a role in flower and fruit development, and that gene expression may be regulated by hormone or environmental conditions. Other regulation levels were also predicted, such as alternative splicing and post-translational modifications. The overall data indicated that citrus HVA22 may be involved in vesicular traffic in stressed cells, and that *CcHVA22d* could be involved in dehydration tolerance.

Keywords: *in silico* analysis; transmembrane domains; *cis*-elements; gene expression; drought; transgenic tobacco

1. Introduction

The *HVA22* gene has been isolated for the first time from the aleurone layer of barley (*Hordeum vulgare*), giving its name [1]. *HVA22* is present in several eukaryotes, such as in plants, fungi and humans, but has not been described in prokaryotes [2]. In plants, *HVA22* is expressed in different tissues, such as seeds, shoots and roots and is induced under several environmental stress conditions (e.g. cold, salinity and drought) mediated by abscisic acid (ABA) [2]. *HVA22* from barley is homologous to *Yop1p* in *Saccharomyces cerevisiae*, *AtHVA22* in *Arabidopsis thaliana*, *AtpCHVA22* in *Aloe vera* L., *OsHVA22* in *Oryza sativa* and *TB2/DP1* in human [3-6]. Approximately 355 *HVA22* homologues were described in eukaryotes. In all homologous, the TB2/DP1 region present in the HVA22 protein structure is conserved [5, 7].

In Arabidopsis, five homologs of *HVA22* were identified, designated as *AtHVA22a*, *AtHVA22b*, *AtHVA22c*, *AtHVA22d* and *AtHVA22e*. These genes showed differential expression in plant tissues, when treated with ABA and subjected to stress conditions [8]. Studies of Yop1p showed that this gene may be involved in vesicular trafficking and transport of substances from the endoplasmic reticulum (ER) to the Golgi complex in yeast [3, 9]. Yop1p/DP1 proteins interact with the reticulon protein Rtn4/NogoA, modulating *in vivo* the protein interaction and ER functions [10, 11]. According to the environmental conditions, the lipid composition of the plasma membrane in yeast and plants could be modified by transport vesicles that deliver the newly synthesized lipids to the membrane [12]. However, the role of HVA22 in this process is still unknown and, more generally, the *HVA22* gene family has been

a poorly studied, even if its potential in response to abiotic stresses has already been highlighted [8].

To our knowledge, in citrus plants, neither genome-wide characterization of the *HVA22* family nor functional studies has yet been performed. The availability of the data from the recent sequencing of the genome of some citrus species (https://www.citrusgenomedb.org/) allowed for the genome-wide analysis of gene families as a pre-requisite for functional and/or pre-breeding studies. The present study aimed to characterize the family of *HVA22* in mandarin (*Citrus clementina*) and sweet orange (*Citrus sinensis*) at nucleotide and protein levels, including promoter analysis and expression pattern. It also provides specific information about the *CcHVA22d* gene; transgenic tobacco plant expressing this gene showed higher dehydration tolerance and less oxidative stress than non-transgenic ones.

2. Material and methods

2.1 *In silico* analysis of HVA22 citrus genes and proteins

The identification and structural analysis of the *HVA22* genes (introns/exons) were performed using the Citrus Genome Database (https://www.citrusgenomedb.org/). Open reading frame (ORF) analysis was performed using the ORFinder software (http://www.ncbi.nlm.nih.gov/orffinder/). The isoelectric point (pI) and molecular weight (MW) were predicted using the pI/Mw tool (www.expasy.org). The predictions of the subcellular location of the protein and of the location of the cleavage site were performed by the TargetP 1.1 server (https://www.cbs.dtu.dk/services/TargetP). Analysis of the conserved family and domains was performed using Pfam (http://pfam.sanger.ac.uk/search/sequence) and ProScan software [13]. Transmembrane helices were predicted using TMPred software [14]. Hydropathicity level was identified using the Protparam program

(https://web.expasy.org/protparam/). The putative phosphorylation sites (Ser / Thr / Tyr) and N-glycosylation sites (tipo Asn-X-Ser / Thr) were identified through the NetPhos 3.1 Server [15] and the NetNGlyc 1.0 server (http://www.cbs. dtu.dk/services/NetNGlyc/) respectively. The protein motif analysis was conducted using the program MEME/MAST [16]. The maximum number of motifs was set to 20, the maximum motif length was set to 80 amino acids, the optimum motif width was constrained to between 6 and 300 residues, and the other parameters were used as default.

2.2 Analysis of the promoter regions and chromosomal location of *HVA22* genes

The identification of *cis*-elements in the promoter regions (1500 bp upstream region from the translation start site) was performed using the plantCARE software (sphinx.rug.ac.be:8080/PlantCARE/cgi/index.html) [17]. The chromosomal location of the *HVA22* genes was obtained by screening the GFF3 file of the genome of *C. clementina* and *C. sinensis* deposited in the citrus genome database using the HVA sequence ID.

2.3 Phylogeny

Phylogenetic analysis was performed based on the alignment of the amino acid sequence of the HVA22 proteins from *C. sinensis* and *C. clementina* with HVA22 proteins from *Arabidopsis thaliana*. The sequences were aligned with ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) [18]. The MEGA 6 program was used to construct a phylogenetic tree by using the neighbor-joining statistical method [19] reliably established by 1000 bootstrap samples.

2.4 *In silico C. sinensis HVA22* gene expression

The expression of the genes of the *C. sinensis HVA22* gene in different tissues (callus, leaf, flower and fruit) was analyzed using the RNA-seq expression data available in the *Citrus sinensis* Annotation Project database version 2.1 (CAP; http://citrus.hzau.edu.cn/orange/) [20]. The Heatmap function of the ComplexHeatmap library from the *R* program [21] was used for the generation and visualization of the transcription level Heatmap.

2.5 Generation of *CcHVA22d* transgenic tobacco plants

The coding sequence of *CcHVA22d* was obtained by RT-PCR from leaves of *C. clementina* cultivated at Embrapa CNPMF (Cruz das Almas, Bahia, Brazil (12°40′39″S latitude, 39°06′23″W longitude and 225m elevation) and submitted to drought as previously described [22]. Briefly, total RNA was extracted using the RNAqueous kit according to the manufacturer instructions (Ambion, Inc.), cDNA was obtained using the Super Script Double Stranded cDNA Synthesis kit according to the manufacturer instructions (Invitrogen), and the *CcHVA22d* cDNA was amplified using specific forward (5'-CGATTCATCACTTTATTCATCAGTCAG-3') and reverse (5'-CTTCTTTTCTTGTAGTTGGTTAGTATGC-3) primers. *CcHVA22d* cDNA was cloned into the pCAMBIA2301 vector (Cambia) under the control of the Cauliflower mosaic virus (CaMV) 35S promoter (*35S-P*) and upstream the terminator (*35S-T*) sequences. The resulting vector contained the *35S-P::CcHVA22d::35S-T* expression cassette as well as the neomycin phosphotransferase II (*nptII*) selectable marker gene and the beta-glucuronidase (GUS) *uidA* as reporter gene. During cloning, the *CcHVA22d* sequence was confirmed by sequencing on the ABI3130XL sequencer according to the Genetic Analyzer User Guide (Applied Biosystems). The *Agrobacterium tumefaciens* strain EHA 105 containing the *pCambiaCcHVA22d* vector was used to transform tobacco plants (*Nicotiana tabacum* cv. Havana) as previously described [23]. Several independent transgenic lines (T0 generation) derived from distinct transformation events were transferred to soil and grown under

125 standardized greenhouse conditions [23]. GUS histochemical assays [24] and PCR screening
126 using the *CcHVA22d* primers indicated above [23] were used to discriminate between
127 transgenic and non-transgenic (control) plants.
128
129 **2.6 Analysis of relative dehydration rate of transgenic *HVA22d* plants**
130 The relative dehydration rate was evaluated in control and transgenic (T_0) tobacco plants by
131 measuring the difference between the initial and final fresh weight (FW) of the leaf tissues in
132 each evaluated period. For this purpose, uniform samples of leaf discs were extracted from the
133 youngest, fully expanded leaf of adult plants, using a 1.0 cm diameter perforator, avoiding the
134 central vein. Then, the leaf discs were submitted to dehydration at room temperature (~22ºC)
135 for 240 min, and fresh weight measurements were performed every 30 min. Each treatment
136 consisted of three replicates with six leaf discs per replicate. Data were submitted to variance
137 analysis and compared by Student's *t*-test with a probability of 1%.
138
139 **2.7 Detection of H_2O_2 in transgenic *HVA22d* plants**
140 To evaluate the oxidative stress in foliar tissues submitted to dehydration, the (control)
141 and transgenic tobacco leaf discs (T_0) were submitted to DAB-HCl (3,3'-diaminobenzidene)
142 histochemical staining assay as previously described [25]. The samples were subjected to
143 vacuum infiltration with 1 mg/ml of DAB-HCl for 4 h. After this period, the leaf discs were
144 boiled in 96% ethanol for 4 h and subsequently photographed in a magnifying glass to detect
145 the brown coloration.
146
147 **3. Results**
148 **3.1 The *HVA22* family in sweet orange and mandarin**

The annotation of the citrus genome database allowed the identification of 12 *HVA22* genes, 6 in *C. clementina* (*CcHVA22a*, *CcHVA22b*, *CcHVA22c*, *CcHVA22d*, *CcHVA22e* and *CcHVA22f*) and 6 in *C. sinensis* (*CsHVA22a*, *CsHVA22b*, *CsHVA22c*, *CsHVA22d*, *CsHVA22e* and *CsHVA22f*) (Table 1). The *C. clementina HVA22* genes were located on chromosomes 3 (*CcHVA22c*), 4 (*CcHVA22b, CcHVA22d* and *CcHVA22e*), 8 (*CcHVA22a*) and 9 (*CcHVA22f*) (Table 1). In *C. sinensis*, the *HVA22* genes were located on chromosomes 5 (*CsHVA22c*), 7 (*CsHVA22b* and *CsHVA22e*) and 8 (*CsHVA22a*) (Table 1). Colinearity analysis showed duplications events between *CcHVA22c* (chromosome 3) and *CcHVA22b* (chromosome 4) (Supplementary material 1) and between *CsHVA22c* (chromosome 5) and *CsHVA22b* (chromosome 7) (data not shown). The *C. clementina* and *C. sinensis* ORFs ranged from 235 to 804 bp, and from 285 to 570 bp, respectively (Table 1; Supplementary material 2). The exon quantity in *C. clementina* and *C. sinensis* varied from 2 to 8, and the intron quantity from 1 to 7 (Table 1; Figure 1).

3.2 Promoter sequence analysis of the citrus *HVA22* genes

A 1500 bp fragment belonging to the upstream region of each *HVA* gene was analyzed to find plant specific *cis*-elements using the PlantCARE database (Supplementary material 3). The TATA and CAAT-box elements were found in all citrus *HVA22* promoter regions; the other *cis*-elements varied between sequence promoters (Fig. 2, Supplementary material 4). Light-responsive *cis*-elements (BOX 4, G-BOX and sp1 elements) and drought-induced element (MYB) were found in great amounts in all promoter sequences (Fig. 2, Supplementary material 4). In smaller proportion, *cis*-elements were found that were responsive to i) hormone or inducers such as auxin, gibberellin, salicylic acid, methyljasmonate, and abscisic acid; and ii) biotic or abiotic stresses such as elicitors, anaerobiosis, defense and stress, and anoxia (Fig. 2, Supplementary material 4). Other *cis*-

elements related to plant development such as seed-specific regulation, meristem expression, zein metabolism, circadian control, endosperm expression were also present in the promoters of the citrus *HVA* genes (Fig. 2, Supplementary material 4).

3.3 Analysis of the citrus HVA proteins

The amino acid amount of HVA22 proteins ranged from 94 to 267 in *C. clementina* and from 94 to 189 in *C. sinensis* (Table 2, Supplementary material 5). Most of the citrus HVA22 presented basic pI, except CcHVA22a, CcHVA22c, CcHVA22c*, CsHVA22a, and CsHVA22c (Table 2). The citrus HVA22 proteins showed transmembrane location with mainly 2-3 transmembrane domains (Table 2). The hydropathicity ranged from -0.272 to 0.419 in *C. clementina* and from -0.335 to 0.315 in *C. sinensis* (Table 2). Citrus HVA22 molecular weight (MW) ranged from 11.512 (CcHVA22f) to 31.072 kDa (CcHVA22b; Table 2). The conserved TB2/DP1 region was found in all the citrus HVA22 (Fig. 3; Supplementary material 5). Putative phosphorylation sites were found in all citrus HVA22 proteins, and N-glycosylation site were found in most of them (Table 3, Supplementary material 5). Most of the citrus HVA22 proteins also contained casein kinase phosphorylation sites; these sites were spread along the protein sequence and ranged from one to five (Supplementary material 5). The protein sequence identity varied from 30% to 93% between CcHVA22 proteins and from 25% to 94% in CsHVA22 proteins (Supplementary material 6). The greatest degree of identity (100%) was observed between CcHVA22a and CsHVA22a, CcHVA22b and CsHVA22b, CcHVA22c/CcHVA22c* and CsHVA22c, CcHVA22d and CsHVA22d (Supplementary material 6). The percentage of identity between the two proteins resulting from the alternative transcripts of the gene was i) 93% for CcHVA22c; ii) 94% for CsHVA22a; iii) 95% for CsHVA22b; and iv) 100% of CsHVA22d (Supplementary material 6).

3.4 Phylogenetic relationships of the *HVA22* genes of citrus and Arabidopsis

Phylogenetic analysis of the HVA22 sequences of citrus and *A. thaliana* showed that CcHVA22d, CsHVA22d and CsHVA22d* were clustered with the sequence AtHVA22d and AtHVA22e, while CcHVA22c, CcHVA22c* and CsHVA22c were clustered with AtHVA22a (Fig. 4). The other citrus HVA22 sequences did not clustered with *A. thaliana* sequences. CcHVA22e, CsHVA22e, CcHVA22f and CsHVA22f were grouped in the same cluster, as well as CsHVA22b, CsHVA22b* and CcHVA22b, and CcHVA22a, CsHVA22a and CsHVA22a* (Fig. 4).

3.5 *In silico CsHVA22* gene expression

The expression of the *CsHVA22* genes was previously obtained and was available in the CAP database. Four tissues were analyzed: callus, flower, leaf and fruit. The *CsHVA22d* gene showed high expression levels in the fruit (> 15 Reads Per Kilobase Million / RPKM) and callus (17.54 RPKM) and low expression levels in flowers and leaves (1.64 and 1.57 RPKM respectively) (Fig. 5, Supplementary material 7). *CsHVA22a* and *CsHVA22c* showed a similar pattern of expression in different tissues with a general quite high expression (>10 RPKM for *CsHVA22a*; > 10 RPKM in fruits > 8 RPKM in flower > 7RPKM in callus and leaf for *CsHVA22c*) (Fig. 5, Supplementary material 7). *CsHVA22e* gene showed highest expression in flower (6.37 RPKM), followed by leaf (4.15 RPKM) and fruit (1.09 RPKM), in callus this expression was very low (<1 RPKM) (Fig. 5, Supplementary material 7). The *CsHVA22b* expression did not exceed 5 RPKM; the lowest level was observed in fruit (0.24 RPKM) (Fig. 5, Supplementary material 7). *CsHVA22b* and *CsHVA22e* showed similar expression pattern and were clustered in the heatmap analysis (Fig. 5). No expression was detected for *CsHVA22f* (data not shown).

2.6 Dehydration analysis and H_2O_2 detection in *CcHVA22d* transgenic tobacco plants

Genetic transformation of tobacco leaf explants was made by *Agrobacterium tumefaciens* co-cultivation method using the *35S::CcHVA22d::pCAMBIA2301* construct. After regeneration, the T_0 transformed lines were confirmed by GUS histochemical test and PCR; five transgenic lines were selected according to both confirmations (Fig. 6A, Supplementary material 8). The L2 and L5 lines showed a more intense blue coloration corresponding to GUS test, followed by the L1, and then the L3 and L4 lines (Fig. 6A). All transgenic lines showed a 534 bp band on electrophoresis gel corresponding to the gene amplification (Supplementary material 8). Transgenic *CcHVA22d* tobacco plants were evaluated under short-term dehydration conditions (Fig. 6B). All transgenic lines lost significantly less water than WT from 30 min to 240 min after leaf disc dehydration (Fig. 6B). Among the transgenic lines, L3 and L5 showed the lowest dehydration rates (Fig. 6B). The dehydration tolerance of transgenic tobacco lines were correlated to oxidative stress levels by detecting H_2O_2 in the same leaf discs using the DAB-HCl histochemical test (Fig. 6C). The brown coloration indicator of H_2O_2 presence was less intense in the transgenic line leaf discs than in the WT ones (Fig. 6C).

4. Discussion

Here, we analyzed the HVA family from citrus. The gene amount was the same in both *C. clementina* and *C. sinensis* (6 genes; Fig. 1, Table 1) and close to the number of *HVA22* genes observed in Arabidopsis (5 genes; [8]). Interestingly, the 100% of identity was observed between the same ID genes from both *C. clementina* and *C. sinensis* (e.g. CcHVA22a and CsHVA22a; Supplementary material 6). These sequence pairs from both citrus species, also clustered together in the phylogenetic analysis (Fig. 4), indicating they

could be considered as orthologues. Only few *HVA22* gene duplication events were observed in citrus (CcHVA22c/CcHVA22b in chromosomes 3/4 and CsHVA22c/CsHVA22b in chromosome 5/7; Supplementary material 1). In Arabidopsis, *AtHVA22d* were located on chromosome 4 and *AtHVA22e* on chromosome 5, and evidences indicated that these two genes arose from a same ancestral gene via duplication events and reorganization [8, 26, 27]. Gene duplication could be considered as the primary source of new gene generation, participating of the genome complexity increase. It also contributed to evolution of novel gene function associated, among others, to floral structure production or stress adaptation [28]. In addition to sequence variations, the analysis of the gene duplication and the evolution of paralogs need to take in consideration other characteristics such as gene expression [27]. Divergent expression patterns observed for some duplicate genes could be due to subfunctionalization and/or neofunctionalization [28]. Here, *CsHVA22b* and *CsHVA22c* showed similar expression pattern (expression in all tissues) but with a lower intensity for the *CsHVA22b* gene, mainly in fruits in which the expression is almost null (Fig. 5). This result could be consistent with subfunctionalization model via duplication–degeneration–complementation processes, in which one of the paralogs (*CsHVA22b*) loss a subfunction of the original gene (degeneration; loss of function in fruit). In this model, both duplicates have been kept to maintain the ancestral functionality (complementation) [27, 28]. Moreover, the complete loss-of-function and/or the absence of gene expression (e.g. *CsHVA22f*, data not shown; *AtHVA22f* [8]) could be associated to pseudogenization, i.e. presence of undeleted nonfunctional duplicates [28]. On the other hand, it has been shown that whole-genome duplication has contributed to important agronomic traits in several crops [28], and for this reason the impacts of gene duplication will be important to future plant studies mainly the agronomically important ones. Interestingly, the duplicated genes in citrus corresponded to the orthologues between *C. clementina* and *C. sinensis* (CcHVA22c/CsHVA22c and

CcHVA22b/CsHVA22b) (Supplementary material 1 and 6). Considering that *C. clementina* is descendant from *C. sinensis*, it could be suggested that these duplication events occurred upstream this crossing via the sweet orange parental [29]. The identification of sequence variations (e.g. polymorphism) in orthologs and/or paralogs could be usefully used in intra- or interspecific diversity studies and/or breeding programs of citrus [29-33].

The citrus HVA22 genes presented alternative transcription; thus 12 citrus genes encoded 16 proteins (Fig. 1, Table 1, Table 2). In general the amount of introns was low among the homologues and this may indicate a shorter period to respond to stimuli caused by a stress condition [34]. This small amount of introns has already been observed for other citrus genes that respond to abiotic stresses [35]. It has been reported that the majority of alternatively spliced RNAs bind to ribosomes and consequently are translated [36]. Moreover, most of the alternative proteins did not present neither disrupted structures nor functions [36]. Here, the percentage of identity between the two proteins resulting from the alternative transcripts of the gene was high (from 93 to 100%; Supplementary material 6), suggesting that the alternative proteins may have real biological function. At the protein level, the citrus HVA22 sequences showed an average of 156 amino acids, close to the 130 amino acids from barley HVA22 [1, 2]. The barley HVA22 protein showed a pI of 9.2 [1, 2], close to the one of most of the citrus HVA22 proteins. The average of the citrus HVA22 MW was about 18 kDa, close to the MW of the HVA22 from *Aloe vera* (15 kDa; [5]). Some of the citrus HVA22 proteins contained casein kinase II phosphorylation sites (Supplementary material 5). Analysis of the amino acid sequence of HVA22 in barley demonstrated that this protein also contained two of casein kinase II phosphorylation sites (SKVD sequence), one between region 36 and 39 and another between region 54 and 59, the latter being conserved in several homologues [1, 2]. Studies have shown that CK2 has action on several cellular functions, including maintenance of cell viability [2]. The protein phosphorylation mediated by kinases

and phosphatases plays essential roles in plant response to many stresses and environmental signals (such as drought) and are convenient for the rapid regulation of protein activity in various signaling pathways [37].

All the citrus HVA22 proteins presented transmembrane location due to several transmembrane domains (Table 2) located in the TB2/DP1 region (Fig. 3; Supplementary material 5). TB2-like 1 is a membrane protein belonging to the YOP1/TB2/DP1/HVA22 family that may play an important role in the intracellular membrane trafficking in retinal ganglion cells from animals [38]. The TB2/DP1 region from barley and its orthologue in yeast, Yop1p, were integral proteins of the endoplasmic reticulum membrane (ER), inserted into the lipid bilayer whose function is to maintain its structure [6]. It has been shown that this transmembrane region played an important physiological role in barley but also in yeast and human [6, 10, 38]. In barley aleurone cell, HVA22 negatively regulated vacuolization and programmed cell death, inhibited vesicular trafficking which promotes membrane turnover in stressed cells, nutrient mobilization in seed germination and seedling growth regulation, as well as cell wall components supplying in plant expanding regions (e.g. roots) allowing a greater absorption of water and minerals [7, 10, 39, 40]. In Arabidopsis, the analysis of the three transmembrane domains (TM1, TM2 and TM3) of AtHVA22d by fusion of the truncated domains with GFP (DTM1:GFP, DTM2:GFP, and DTM3:GFP) showed that TM2 was important for protein stability and correct subcellular localization of the protein *in vivo* [7]. Even if the expression and localization of DTM1:GFP and DTM3:GFP were similar to those of the full length protein (AtHVA22d:GFP), the authors did not excluded the possibility that these two domains are also required for the protein function [7]. Transgenic Arabidopsis containing HVA22d silenced by RNAi showed reduction of flower, pollen and seed development associated to enhanced autophagy (e.g. accumulation of vacuoles and cytoplasm invagination) [41]. Similar result was observed in yeast and it has been suggested that HVA22

homologs play a negative role in autophagy, and may be part of a protective mechanism used to suppress unnecessary catabolism while anabolic activities are reduced, as observed in plant submitted to environmental stresses (e.g. high salinity, drought, low temperature) [41].

Citrus sinensis HVA22 gene expression showed that *CsHVA22a*, *CsHVA22c* and *CsHVA22d* presented a high expression compared to the other genes (Fig. 5). The *CsHVA22d* gene was mainly expressed in fruits and callus (Fig. 5). In Arabidopsis, the *HVA22* genes presented differential expression pattern according to plant tissues and to biotic stresses (e.g. ABA, cold stress, dehydration and salt stress) [8]. The *AtHVA22a* gene, which is phylogenetically close to *CsHVA22c/CcHVA22c* (Fig. 4), showed a constitutive high expression in all stress conditions as well as in most of the plant tissues analyzed (rosette leaves, leaves, inflorescence stems and flower buds) [8]. Similar expression pattern was observed for *CsHVA22c* (Fig. 5). The *AtHVA22b* expression is specifically induced by ABA and suppressed by cold stress [8], but in our analysis, this sequence did not clustered to any citrus sequence (Fig. 4). The *AtHVA22c* gene showed a constitutive expression in all stress conditions analyzed – similar but lower in intensity than *AtHVA22a* expression – and in stems and flower buds [8]. *AtHVA22c* is phylogenetically isolated from the other genes (Fig. 4); the closest sequence was *CsHVA22b*, which showed a low expression in the different analyzed citrus tissues (Fig. 5). The *AtHVA22d* gene was highly expressed under ABA, cold stress, dehydration and salt stress, as well as in flower buds, while *AtHVA22e* was mainly expressed under cold stress, dehydration and salt stress, and in stems [8]. *AtHVA22d* and *AtHVA22e* were phylogenetically close to *CsHVA22d*/*CcHVA22d* (Fig. 4) and the *CsHVA22d* expression was high in callus and fruit (Fig. 5). The *CsHVA22e*/*CcHVA22e* and *CsHVA22f*/*CcHVA22f* sequences did not clustered to any Arabidopsis sequences in the phylogenetic analysis (Fig. 4); the *CsHVA22e* expression was low in all tissues (similar pattern to the *CsHVA22b* one; Fig. 5) and the *CsHVA22f* expression was not detected (data not shown). The phylogenetic

and expression analyses of the citrus HVA sequences showed a separation of CsHVA22d/CcHVA22d (close to AtHVA22d and AtHVA22e) from the other citrus sequences (Fig. 4 and 5). As discussed above, the *CcHVA22b* and *CcHVA22c* were shown to be duplications from the same ancestral, and this result may be correlated to the phylogenetic separation between the AtHVA22d/AtHVA22e group (also close to barley HVA22) from the AtHVA22a/AtHVA22b/AtHVA22c group, suggesting that these two groups had diverged before monocots/dicots divergence [8].

The regulation of gene expression depends of *cis*-elements present in the promoter sequences of each gene (Fig. 2, Supplementary material 5). The G-BOX element, present in almost all citrus HVA22 promoters, is known to be associated to the response to several environmental factors and/or internal stimuli such as light, anaerobiosis and hormones [42]. In both citrus species, ABA responsiveness elements were found such as the ABRE element containing the ACGT motif, as also observed in barley HVA1 sequence [42, 43]. In barley, the HVA1 promoter contained distal and proximal coupling elements (CEs), which have a synergistic effect on the absolute expression level as well as on ABA induction [42]. These *cis*-elements were specifically recognized by the HvABI5 bZIP transcription factor, which together with another transcription factor, HvVP1, transactivated the promoter [44, 45]. The citrus *HVA22* promoters contained *cis*-elements related to anaerobic induction, anoxic specific inducibility and drought inducibility (Fig. 2, Supplementary materials 4 and 5) that could be related to root function in stress situation such as flooding or, on the opposite, water loss. Interestingly, it had been shown that the homolog of Arabidopsis HVA22 in yeast, Yop1, and the homolog of Arabidopsis root hair defective 3/RHD3 in yeast, SEY1, presented synthetic enhancement, and could be involved in vesicular traffic regulation in root stressed cells [39]. In barley and Arabidopsis, HVA22 genes showed differential expression under drought stress [2, 8], and here we showed that *CcHVA22d* expressed under constitutive promoter in

transgenic tobacco plants conferred them dehydration tolerance (Fig. 6). In rice, the ABRC HVA1- promoter like motif (which contain the ABRE element and a CE) was found in genes related to biotic stress responses, disease resistance and defense such as NPR1 or WRKY transcription factors [46]. In citrus *HVA22* promoters, several *cis*-elements were related to elicitor response or defense (Fig. 2, Supplementary materials 4 and 5), corroborating the idea that the plant response pathways to biotic and abiotic stresses significantly overlap, and that HVA1-like enhancer modules participate to a complex regulatory network signalized by ABA [46]. The citrus *HVA22* promoter also contained *cis*-element related to plant development mainly seed regulation, endosperm expression and zein metabolism (Fig. 2) that may be correlated to the well-known function of HVA barley gene in seed maturation and dormancy [2, 7]. The accumulation of HVA22 mRNA in aleurone layer during the late stage of seed maturation was associated to high levels of endogenous ABA required for dormancy [7]. The citrus promoters also contained *cis*-elements responsive to gibberellin (Fig. 2). In barley, a treatment with gibberellic acid (GA) was responsible for dormancy break associated to decrease of HVA22 transcripts [1, 7]. It has also been shown that GA and ABA regulate vacuolation associated to PCD in barley aleurone cells, and that overexpression of HVA22 inhibited GA-induced PCD [7]. Finally, the citrus *HVA22* promoters contained *cis*-elements responsive to light (Fig. 2, Supplementary materials 4 and 5), similar to those observed in the promoter of the chalcone synthase, which is induced by light / UV light [47].

According to its expression pattern (high expression in fruit and callus; Fig. 5) and to the presence of drought inducible *cis*-elements in its promoter (e.g. MYB; Fig. 2, Supplementary materials 4 and 5), *CcHVA22d* was selected for further analysis of its involvement in drought tolerance using Agrobacterium mediated-plant transformation (Fig. 6). Transgenic *CcHVA22d* tobacco lines showed lower dehydration rate associated to lower accumulation of H_2O_2 than WT plants (Fig. 6B and 6C) suggesting that *CcHVA22d* conferred

a higher tolerance to drought and less suffered oxidative stress. This result also supports the involvement of citrus *HVA22* genes in response to environmental stresses; more detailed phenotypic characterization of *CcHVA22d* transgenic plants is under way by our research team.

5. Conclusion

To our knowledge, this is the first citrus HVA22 family characterization. Six *HVA22* genes were identified in each citrus species analyzed; orthology was observed between the *C. clementina* and *C. sinensis* genes. Phylogenetic and expression analyses of the citrus and Arabidopsis HVA22 sequences revealed that CsHVA22d sequence clustered with AtHVA22d and that both were expressed in reproductive organs (e.g. flowers, fruits). According to the *cis*-elements present in the citrus *HVA22* promoters, the gene expression may be regulated by several internal and external factors. Other regulation levels were also predicted, such as alternative splicing and post-translational modifications including CK2 phosphorylation, which is involved in plant response to environmental stress. The citrus HVA22 proteins were transmembrane and contained YOP1/TB2/DP1/HVA22 region associated to regulation of vesicular traffic in stressed cells. Moreover, the *CcHVA22d* gene showed to be involved in dehydration tolerance and oxidative stress reduction.

Abbreviations

ABA: abscisic acid; ER: endoplasmic reticulum; MW: molecular weight; ORF: Open reading frame; pI: isoelectric point; RPKM: Reads Per Kilobase Million.

Acknowledgments

MDGF and RJSS were supported by the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB). JAC was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). FM received o Productivity grant from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). This work was supported by CNPq Universal project (n°471742/2013-9) coordinated by FM. We thank Abelmon da Silva Gesteira (Embrapa CNPMF, Brazil) for providing *Citrus clementina* leaves. We thank the Núcleo de Biologia Computacional e Gestão de Informacões Biotecnologicas (NBCIGB) from UESC for providing the infrastructure for the bioinformatics analysis. This work was developed in the frame of the International Consortium in Advance Biology (CIBA; https://www.ciba-network.org/).

Authors' contribution

MDGF executed all the analyses; MDGF, JAC and FM designed the analyses; MDGF, JAC and FM wrote the manuscript; JAC participated to the transgenic plant generation; RJSS participated of the phylogeny analysis; FM was responsible for the financial support of the research and for the advising of MDGF, JAC and RJSS.

References

[1] Q. Shen, S.J. Uknes, T.H. Ho, Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene, Journal of Biological Chemistry, 268 (1993) 23652-23660.

[2] Q. Shen, C.-N. Chen, A. Brands, S.-M. Pan, D.H. Tuan-Hua, The stress- and abscisic acid-induced barley gene HVA22: developmental regulation and homologues in diverse organisms, Plant Molecular Biology, 45 (2001) 327-340.

[3] X. Yang, H.T. Matern, D. Gallwitz, Specific binding to a novel and essential Golgi membrane protein (Yip1p) functionally links the transport GTPases Ypt1p and Ypt31p, The EMBO Journal, 17 (1998) 4954-4963.

[4] B.-Z. Xiao, X. Chen, C.-B. Xiang, N. Tang, Q.-F. Zhang, L.-Z. Xiong, Evaluation of seven function-known candidate genes for their effects on improving drought resistance of transgenic rice under field conditions, Molecular Plant, 2 (2009) 73-83.

[5] K. Sharon, S. Suvarna, Cloning of HVA22 Homolog from Aloe vera and Preliminary Study of Transgenic Plant Development, Int. J. Pure App. Biosci., 5 (2017) 1113-1121.

[6] M. Calero, G.R. Whittaker, R.N. Collins, Yop1p, the Yeast Homolog of the Polyposis Locus Protein 1, Interacts with Yip1p and Negatively Regulates Cell Growth, Journal of Biological Chemistry, 276 (2001) 12100-12112.

[7] W.-J. Guo, T.-H. David Ho, An Abscisic Acid-Induced Protein, HVA22, Inhibits Gibberellin-Mediated Programmed Cell Death in Cereal Aleurone Cells, Plant Physiology, 147 (2008) 1710-1722.

[8] C.-N. Chen, C.-C. Chu, R. Zentella, S.-M. Pan, T.-H. David Ho, AtHVA22 gene family in Arabidopsis: phylogenetic relationship, ABA and stress regulation, and tissue-specific expression, Plant Molecular Biology, 49 (2002) 631-642.

[9] A. De Antoni, J. Schmitzová, H.-H. Trepte, D. Gallwitz, Š. Albert, Significance of GTP Hydrolysis in Ypt1p-regulated Endoplasmic Reticulum to Golgi Transport Revealed by the Analysis of Two Novel Ypt1-GAPs, Journal of Biological Chemistry, 277 (2002) 41023-41031.

[10] G.K. Voeltz, W.A. Prinz, Y. Shibata, J.M. Rist, T.A. Rapoport, A Class of Membrane Proteins Shaping the Tubular Endoplasmic Reticulum, Cell, 124 (2006) 573-586.

[11] J. Hu, Y. Shibata, C. Voss, T. Shemesh, Z. Li, M. Coughlin, M.M. Kozlov, T.A. Rapoport, W.A. Prinz, Membrane Proteins of the Endoplasmic Reticulum Induce High-Curvature Tubules, Science, 319 (2008) 1247-1250.

[12] I. Hamrouni, H.B. Salah, B. Marzouk, Effects of water-deficit on lipids of safflower aerial parts, Phytochemistry, 58 (2001) 277-280.

[13] E. Quevillon, V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, R. Lopez, InterProScan: protein domains identifier, Nucleic Acids Research, 33 (2005) W116-W120.

[14] Hofmann, TMbase - A database of membrane spanning proteins segments, Biol. Chem. Hoppe-Seyler, 374 (1993).

[15] N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, J Mol Biol, 294 (1999) 1351-1362.

[16] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME Suite: tools for motif discovery and searching, Nucleic Acids Res, 37 (2009) W202-W208.

[17] M. Lescot, P. Déhais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouzé, S. Rombauts, PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Research, 30 (2002) 325-327.

[18] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research, 22 (1994) 4673-4680.

[19] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol Biol Evol, 4 (1987) 406-425.

[20] J. Wang, D. Chen, Y. Lei, J.-W. Chang, B.-H. Hao, F. Xing, S. Li, Q. Xu, X.-X. Deng, L.-L. Chen, Citrus sinensis Annotation Project (CAP): A Comprehensive Database for Sweet Orange Genome, PLoS ONE, 9 (2014) e87723.

[21] Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics, 32 (2016) 2847-2849.

[22] D.M. Neves, D.D.S. Santana-Vieira, M.S. Dória, L. Freschi, C.F. Ferreira, W.d.S. Soares Filho, M.G.C. Costa, M.A. Coelho Filho, F. Micheli, A.d.S. Gesteira, Recurrent water deficit causes alterations in the profile of redox proteins in citrus plants, Plant Physiology and Biochemistry, 132 (2018) 497-507.

[23] L.C. Cidade, T.M. de Oliveira, A.F.S. Mendes, A.F. Macedo, E.I.S. Floh, A.S. Gesteira, W.S. Soares-Filho, M.G.C. Costa, Ectopic expression of a fruit phytoene synthase from Citrus paradisi Macf. promotes abiotic stress tolerance in transgenic tobacco, Molecular Biology Reports, 39 (2012) 10201-10209.

[24] R.A. Jefferson, T.A. Kavanagh, M.W. Bevan, GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants, The EMBO Journal, 6 (1987) 3901-3907.
[25] H. Thordal-Christensen, Z. Zhang, Y. Wei, D.B. Collinge, Subcellular localization of H2O2 in plants. H2O2 accumulation in papillae and hypersensitive response during the barley—powdery mildew interaction, The Plant Journal, 11 (1997) 1187-1194.
[26] G. Blanc, A. Barakat, R. Guyot, R. Cooke, M. Delseny, Extensive Duplication and Reshuffling in the Arabidopsis Genome, The Plant Cell, 12 (2000) 1093-1101.
[27] C.-N. Chen, Y.-C. Chiang, T.-H. David Ho, B.A. Schaal, T.-Y. Chiang, Coalescent processes and relaxation of selective constraints leading to contrasting genetic diversity at paralogs AtHVA22d and AtHVA22e in Arabidopsis thaliana, Molecular Phylogenetics and Evolution, 32 (2004) 616-626.
[28] N. Panchy, M. Lehti-Shiu, S.-H. Shiu, Evolution of Gene Duplication in Plants, Plant Physiology, 171 (2016) 2294-2316.
[29] G.A. Wu, J. Terol, V. Ibanez, A. López-García, E. Pérez-Román, C. Borredá, C. Domingo, F.R. Tadeo, J. Carbonell-Caballero, R. Alonso, F. Curk, D. Du, P. Ollitrault, M.L. Roose, J. Dopazo, F.G. Gmitter, D.S. Rokhsar, M. Talon, Genomics of the origin and evolution of Citrus, Nature, (2018).
[30] F. Curk, G. Ancillo, F. Ollitrault, X. Perrier, J.-P. Jacquemoud-Collet, A. Garcia-Lor, L. Navarro, P. Ollitrault, Nuclear Species-Diagnostic SNP Markers Mined from 454 Amplicon Sequencing Reveal Admixture Genomic Structure of Modern Citrus Varieties, PLoS ONE, 10 (2015) e0125628.
[31] F. Curk, F. Ollitrault, A. Garcia-Lor, F. Luro, L. Navarro, P. Ollitrault, Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers, Annals of Botany, 117 (2016) 565-583.
[32] G.A. Wu, S. Prochnik, J. Jenkins, J. Salse, U. Hellsten, F. Murat, X. Perrier, M. Ruiz, S. Scalabrin, J. Terol, M.A. Takita, K. Labadie, J. Poulain, A. Couloux, K. Jabbari, F. Cattonaro, C. Del Fabbro, S. Pinosio, A. Zuccolo, J. Chapman, J. Grimwood, F.R. Tadeo, L.H. Estornell, J.V. Muñoz-Sanz, V. Ibanez, A. Herrero-Ortega, P. Aleza, J. Pérez-Pérez, D. Ramón, D. Brunel, F. Luro, C. Chen, W.G. Farmerie, B. Desany, C. Kodira, M. Mohiuddin, T. Harkins, K. Fredrikson, P. Burns, A. Lomsadze, M. Borodovsky, G. Reforgiato, J. Freitas-Astúa, F. Quetier, L. Navarro, M. Roose, P. Wincker, J. Schmutz, M. Morgante, M.A. Machado, M. Talon, O. Jaillon, P. Ollitrault, F. Gmitter, D. Rokhsar, Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication, Nature biotechnology, 32 (2014) 656-662.
[33] D. Ahmed, A. Comte, F. Curk, G. Costantino, F. Luro, A. Dereeper, P. Mournet, Y. Froelicher, P. Ollitrault, Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: a case study in diploid and polyploid citrus, Annals of Botany, (2019).
[34] D.C. Jeffares, C.J. Penkett, J. Bähler, Rapidly regulated genes are intron poor, Trends in Genetics, 24 (2008) 375-378.
[35] J. Araújo Castro, M.D. Gomes Ferreira, R.J. Santana Silva, B.S. Andrade, F. Micheli, Alternative oxidase (AOX) constitutes a small family of proteins in Citrus clementina and Citrus sinensis L. Osb, PLoS ONE, 12 (2017) e0176878-e0176878.
[36] A. Reyes, W. Huber, Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues, Nucleic Acids Research, 46 (2018) 582-592.
[37] X.-Y. Wu, T. Li, A Casein Kinase II Phosphorylation Site in AtYY1 Affects Its Activity, Stability, and Function in the ABA Response, Frontiers in Plant Science, 8 (2017) 323-323.
[38] H. Sato, H. Tomita, T. Nakazawa, S. Wakana, M. Tamai, Deleted in Polyposis 1-like 1 Gene (Dp1l1): A Novel Gene Richly Expressed in Retinal Ganglion Cells, Investigative Ophthalmology & Visual Science, 46 (2005) 791-796.
[39] A. Brands, T.-h.D. Ho, Function of a plant stress-induced gene, HVA22. Synthetic enhancement screen with its yeast homolog reveals its role in vesicular traffic, Plant Physiology, 130 (2002) 1121-1131.
[40] S. Gilroy, D.L. Jones, Through form to function: root hair development and nutrient uptake, Trends in Plant Science, 5 (2000) 56-60.
[41] C.-N.N. Chen, H.-R. Chen, S.-Y. Yeh, G. Vittore, T.-H.D. Ho, Autophagy is enhanced and floral development is impaired in AtHVA22d RNA interference Arabidopsis, Plant Physiology, 149 (2009) 1679-1689.
[42] Q. Shen, P. Zhang, T.H. Ho, Modular nature of abscisic acid (ABA) response complexes: composite promoter units that are necessary and sufficient for ABA induction of gene expression in barley, The Plant Cell, 8 (1996) 1107-1119.
[43] R. Velasco, F. Salamini, D. Bartels, Gene structure and expression analysis of the drought- and abscisic acid-responsive CDeT11-24 gene family from the resurrection plant Craterostigma plantagineum Hochst, Planta, 204 (1998) 459-471.
[44] J. Casaretto, T.-h.D. Ho, The transcription factors HvABI5 and HvVP1 are required for the abscisic acid induction of gene expression in barley aleurone cells, The Plant Cell, 15 (2003) 271-284.
[45] J.A. Casaretto, T.-h.D. Ho, Transcriptional regulation by abscisic acid in barley (Hordeum vulgare L.) seeds involves autoregulation of the transcription factor HvABI5, Plant Molecular Biology, 57 (2005) 21-34.

[46] C. Ross, Q.J. Shen, Computational Prediction and Experimental Verification of HVA1-like Abscisic Acid Responsive Promoters in Rice (Oryza sativa), Plant Molecular Biology, 62 (2006) 233-246.
[47] Q. Shen, T.H. Ho, Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel cis-acting element, The Plant Cell, 7 (1995) 295-307.

Tables

Table 1. Characteristics of the *HVA22* genes present in *C. clementina* and *C. sinensis* genomes. ORF: open reading frame. (*) indicated the gene ID of the alternative transcript of the *HVA22* gene.

Species	Gene name	Gene ID	Location	ORF size (bp)	Quantity of introns	Quantity of exons
C. clementina	*CcHVA22a*	Ciclev10029342m	Chromosome 8	570	7	8
	CcHVA22b	Ciclev10033698m	Chromosome 4	804	6	7
	CcHVA22c	Ciclev10022553m	Chromosome 3	516	4	5
	*CcHVA22c**	Ciclev10022663m	Chromosome 3	462	2	3
	CcHVA22d	Ciclev10033011m	Chromosome 4	435	4	5
	CcHVA22e	Ciclev10032960m	Chromosome 4	477	4	5
	CcHVA22f	Ciclev10006951m	Chromosome 9	235	1	2
C. sinensis	*CsHVA22a*	orange1.1g029734m	Chromosome 8	570	7	8
	*CsHVA22a**	orange1.1g032614m	Chromosome 8	414	5	6
	CsHVA22b	orange1.1g030361m	Chromosome 7	540	4	5
	*CsHVA22b**	orange1.1g031732m	Chromosome 7	465	2	3
	CsHVA22c	orange1.1g030780m	Chromosome 5	516	4	5
	CsHVA22d	orange1.1g032264m	Chromosome -	435	4	5
	*CsHVA22d**	orange1.1g033193m	Chromosome -	378	3	4
	CsHVA22e	orange1.1g033043m	Chromosome 7	387	3	4
	CsHVA22f	orange1.1g047167m	Chromosome -	285	1	2

Table 2. Characteristics of the HVA22 proteins present in the citrus genomes. aa: amino acids; GRAVY: grand average of hydropathicity; Mw: molecular weight; pI: isoelectric point.

Protein name	Protein size (aa)	pI	Mw (kDa)	Subcellular location*	GRAVY
CcHVA22a	189	6.51	21.604	Transmembrane (3)	0.029
CcHVA22b	267	9.68	31.072	Transmembrane (4)	-0.272
CcHVA22c	171	6.29	20.02	Transmembrane (3)	-0.081
CcHVA22c*	153	6.51	18.367	Transmembrane (2)	-0.222
CcHVA22d	144	9.45	16.869	Transmembrane (3)	-0.085
CcHVA22e	158	9.37	18.399	Transmembrane (2)	0.115
CcHVA22f	94	9.01	11.512	Transmembrane (2)	0.419
CsHVA22a	189	6.51	21.604	Transmembrane (2)	0.029
CsHVA22a*	137	8.85	15.79	Transmembrane (3)	0.256
CsHVA22b	179	8.57	21.161	Transmembrane (3)	-0.217
CsHVA22b*	154	8.81	18.511	Transmembrane (2)	-0.335
CsHVA22c	171	6.29	20.033	Transmembrane (3)	-0.081
CsHVA22d	144	9.45	16.869	Transmembrane (3)	-0.085
CsHVA22d*	125	9.33	14.727	Transmembrane (2)	-0.205
CsHVA22e	128	9.57	14.774	Transmembrane (1)	-0.141
CsHVA22f	94	9.61	11.697	Transmembrane (2)	0.315

* Transmembrane domain number is indicated under parenthesis.

Table 3. Post-translational modifications of citrus HVA22 proteins.

Protein	Phosphorylation sites	N-glycosylation sites
CcHVA22a	S_4, S_{44}, S_{64}, S_{95}, S_{105}, S_{124}, S_{130}, S_{134}, S_{179}, S_{183}, S_{185}, T_{38}, T_{96}, T_{145}, T_{157}, T_{167}	N_{155}

CcHVA22b	S_{24}, S_{38}, S_{43}, S_{44}, S_{77}, S_{94}, S_{124}, S_{132}, S_{134}, S_{147}, S_{202}, S_{204}, S_{205}, S_{238}, S_{252}, S_{255}, T_{26}, T_{74}, T_{79}, T_{141}, T_{150}, T_{225}, T_{236}, T_{246}, Y_{197}, Y_{209}, Y_{230}, Y_{258}, Y_{265}	-
CcHVA22c	S_{7}, S_{33}, S_{41}, S_{56}, S_{76}, S_{125}, S_{156}, S_{163}, T_{39}, T_{59}, T_{64}, T_{132}, T_{143}, T_{164}, Y_{31}, Y_{96}, Y_{137}, Y_{160}, Y_{165}, Y_{167}, Y_{171}	N_{141}
CcHVA22c*	S_{15}, S_{23}, S_{38}, S_{58}, S_{107}, S_{138}, S_{145}, T_{21}, T_{41}, T_{46}, T_{114}, T_{125}, T_{146}, Y_{78}, Y_{119}, Y_{142}, Y_{147}, Y_{149}, Y_{153}	N_{112}
CcHVA22d	S_{13}, S_{34}, S_{36}, S_{51}, S_{72}, S_{118}, S_{119}, T_{121}, T_{123}, Y_{144}	-
CcHVA22f	S_{11}, S_{26}, T_{6}, T_{38}, Y_{1}, Y_{15}	-
CcHVA22e	S_{28}, S_{34}, S_{36}, S_{51}, S_{59}, S_{110}, S_{111}, S_{131}, S_{133}, T_{45}, T_{155}, Y_{100}, Y_{106}, Y_{135}	N_{112}
CsHVA22a	S_{4}, S_{44}, S_{64}, S_{95}, S_{105}, S_{124}, S_{130}, S_{134}, S_{179}, S_{183}, S_{185}, T_{38}, T_{96}, T_{145}, T_{157}, T_{167}	N_{155}
CsHVA22a*	S_{4}, S_{44}, S_{64}, S_{95}, S_{105}, S_{124}, T_{38}, T_{96}	-
CsHVA22b	S_{6}, S_{36}, S_{44}, S_{46}, S_{59}, S_{114}, S_{116}, S_{117}, S_{150}, S_{167}, T_{53}, T_{62}, T_{137}, T_{148}, T_{158}, Y_{109}, Y_{142}, Y_{170}, Y_{177}	-
CsHVA22b*	S_{11}, S_{19}, S_{21}, S_{34}, S_{89}, S_{91}, S_{92}, S_{125}, S_{139}, S_{142}, T_{28}, T_{37}, T_{112}, T_{123}, T_{133}, Y_{84}, Y_{96}, Y_{117}, Y_{145}, Y_{152}	-
CsHVA22c	S_{7}, S_{41}, S_{56}, S_{76}, S_{125}, S_{156}, S_{163}, T_{39}, T_{59}, T_{64}, T_{132}, T_{143}, T_{164}, Y_{31}, Y_{96}, Y_{137}, Y_{160}, Y_{165}, Y_{167}, Y_{171}	N_{141}
CsHVA22d	S_{13}, S_{34}, S_{36}, S_{51}, S_{72}, S_{118}, S_{119}, T_{121}, T_{123}, Y_{144}	-
CsHVA22d*	S_{15}, S_{17}, S_{32}, S_{53}, S_{99}, S_{100}, T_{102}, T_{104}, Y_{125}	-
CsHVA22e	S_{19}, S_{25}, S_{27}, S_{78}, S_{80}, S_{81}, S_{101}, S_{103}, T_{28}, T_{36}, T_{125}, Y_{76}, Y_{105}	N_{82}
CsHVA22f	S_{11}, S_{26}, T_{6}, T_{38}, Y_{1}	-

Figure legends

Figure 1. Structure of *HVA22* genes from *C. clementina* and *C. sinensis*. The exons and the introns were represented in orange and in black, respectively. (*) indicated the alternative transcript of the *HVA* gene with the same name from the same species.

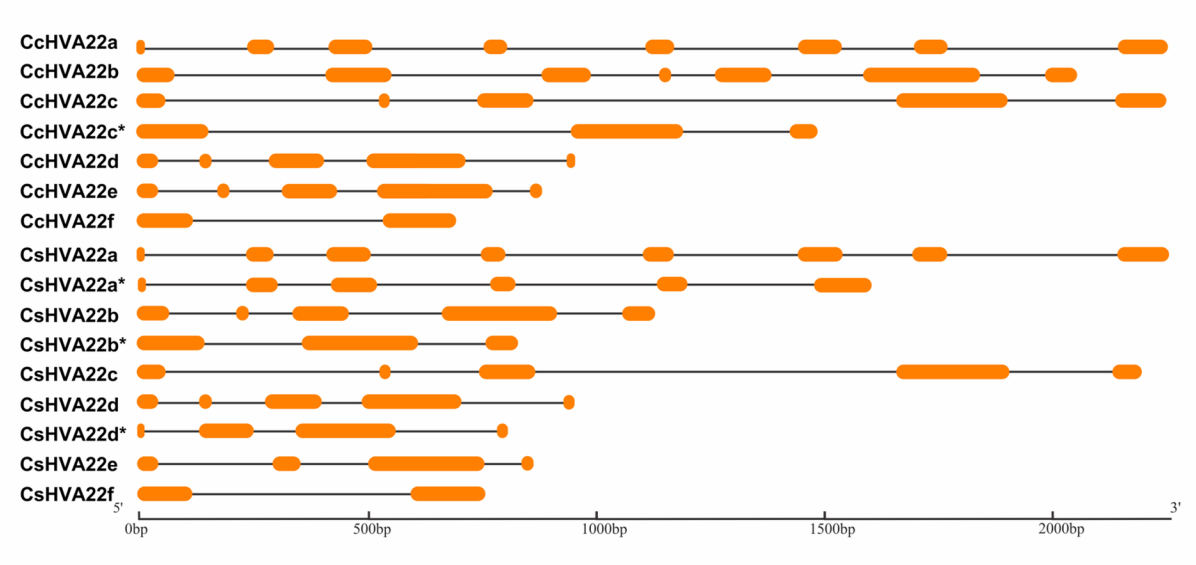
Figure 2. *Cis*-elements present in the promoter region of citrus *HVA* genes. **A)** *Cis*-elements from the analysis of the *C. clementina* promoter. **B)** *Cis*-elements from the analysis of the *C. sinensis* promoter. The *cis*-elements were analyzed in the 1500 bp upstream from the translation start site using the plantCARE database.

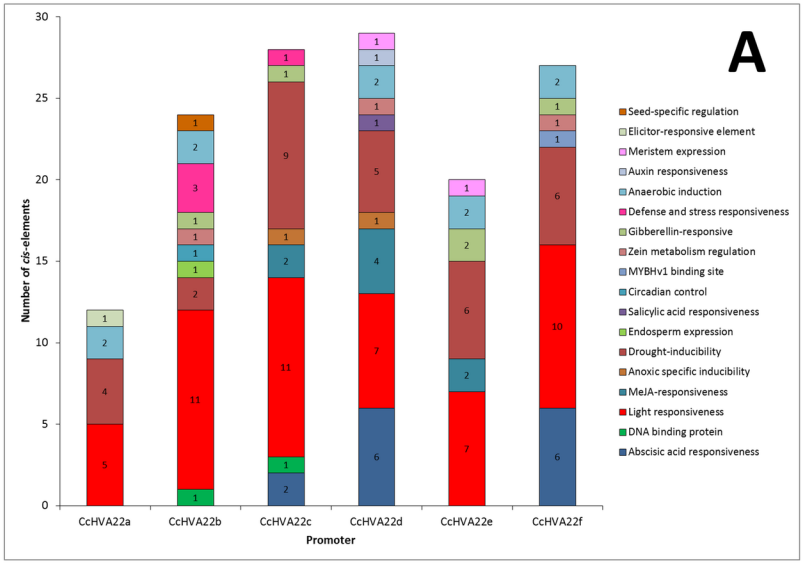
Figure 3. Conserved motifs in *citrus* HVA proteins obtained by the MEME program. Conserved amino acid sequence in the TB2/DP1/HVA22 region.

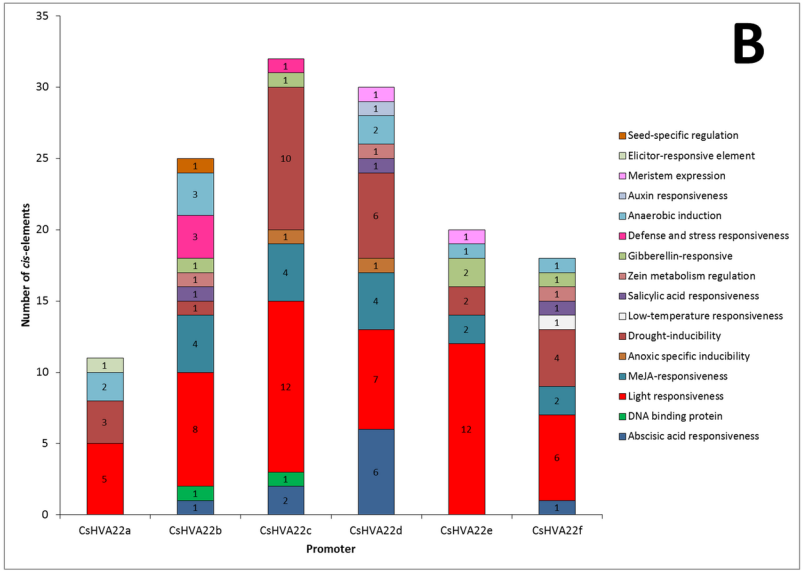
589 **Figure 4**. Phylogenetic tree obtained with the HVA22 proteins of *A. thaliana*, *C. sinensis* and
590 *C. clementina*.

591 **Figure 5.** Heatmap of the *CsHVA22* gene expression in different tissues. Pattern generated
592 through *R* software.

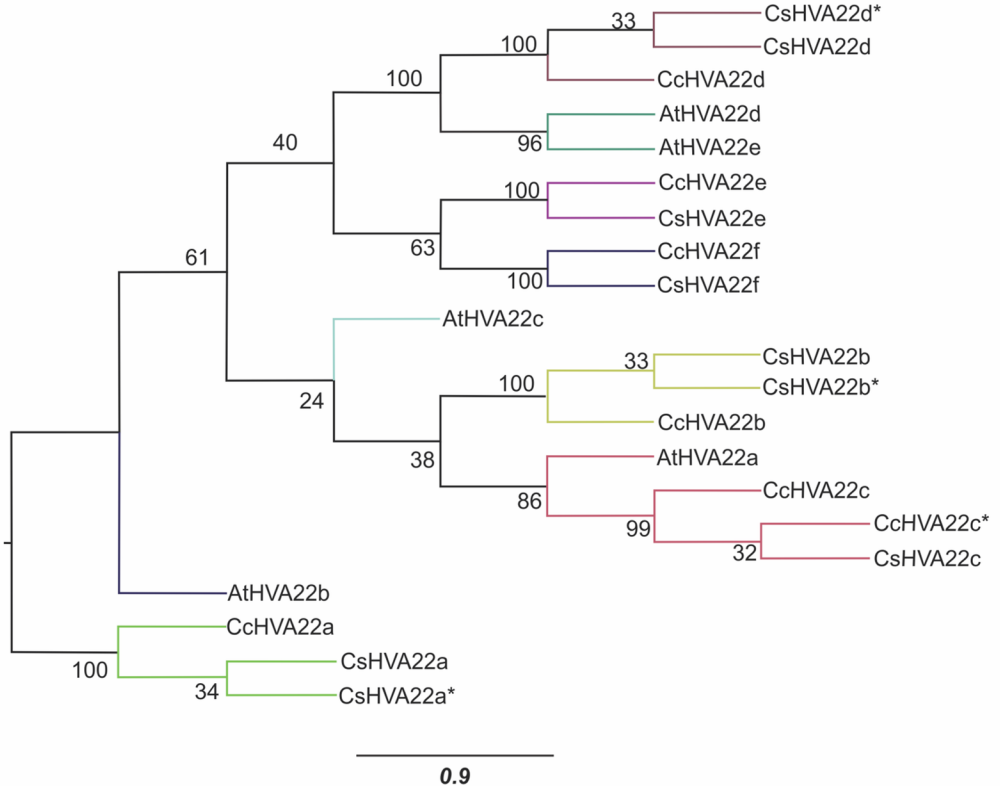
593 **Figure 6.** *CcHVA22d* transgenic tobacco plants analysis. **A.** GUS histochemical test. **B.**
594 Dehydration rate analysis. (*) indicates that all the transgenic lines are statistically different
595 from the WT plant. **C.** H_2O_2 detection using DAB-HCl. L1 to L5: transgenic *CcHVA22d*
596 lines; WT: wild-type (non transgenic plants).

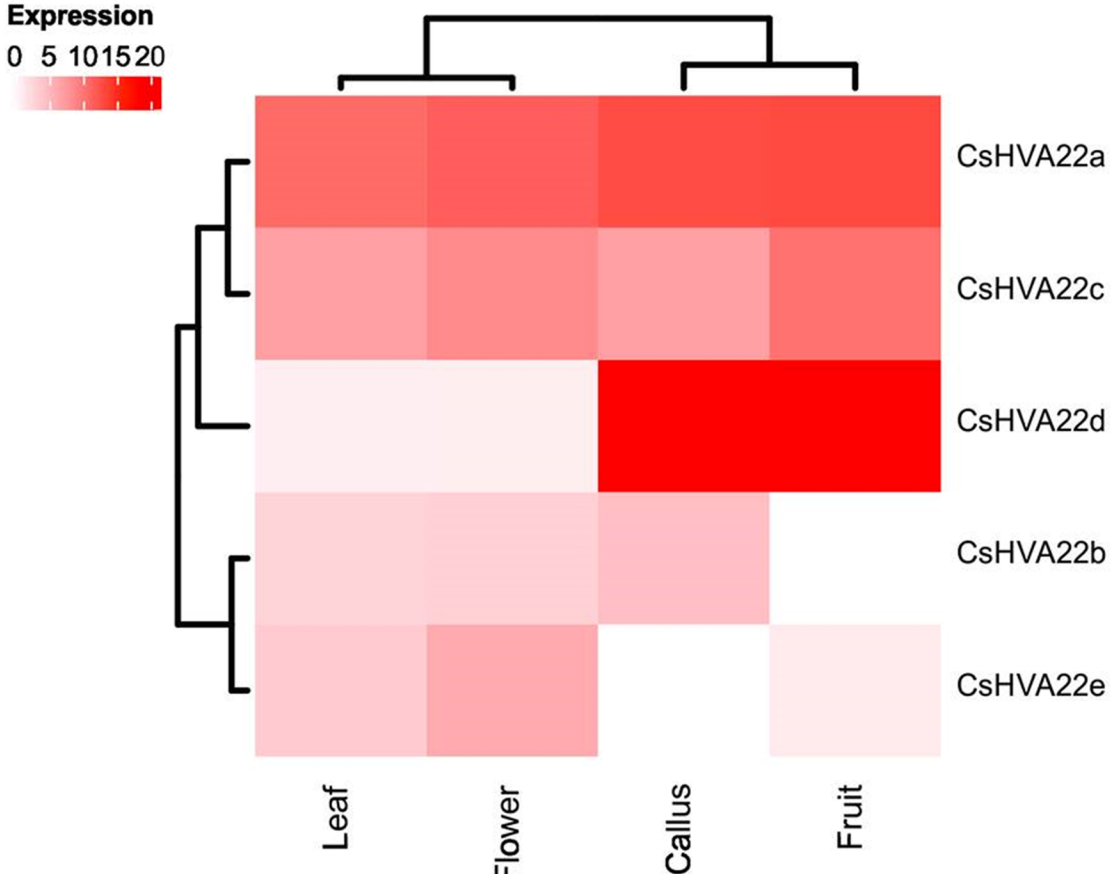

CcHVA22a
CcHVA22b
CcHVA22c
CcHVA22c*
CcHVA22d
CcHVA22e
CcHVA22f
CsHVA22a
CsHVA22a*
CsHVA22b
CsHVA22b*
CsHVA22c
CsHVA22d
CsHVA22d*
CsHVA22e
CsHVA22f
5'
3'
0bp
500bp
1000bp
1500bp
2000bp

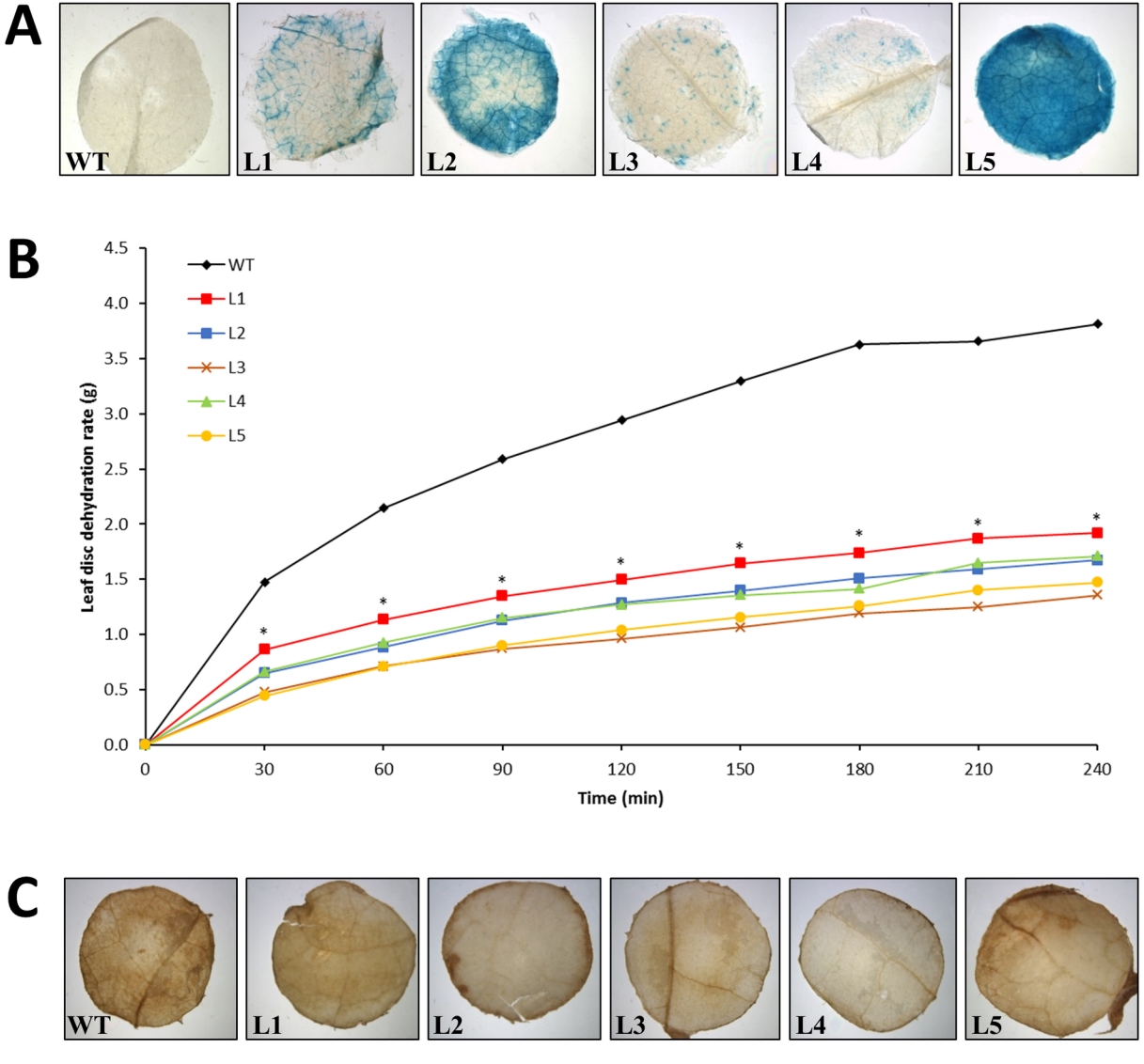

A
Number of cis-elements
Promoter
CcHVA22a
CcHVA22b
CcHVA22c
CcHVA22d
CcHVA22e
CcHVA22f
Seed-specific regulation
Elicitor-responsive element
Meristem expression
Auxin responsiveness
Anaerobic induction
Defense and stress responsiveness
Gibberellin-responsive
Zein metabolism regulation
MYBHv1 binding site
Circadian control
Salicylic acid responsiveness
Endosperm expression
Drought-inducibility
Anoxic specific inducibility
MeJA-responsiveness
Light responsiveness
DNA binding protein
Abscisic acid responsiveness


B
Number of cis-elements
Promoter
CsHVA22a
CsHVA22b
CsHVA22c
CsHVA22d
CsHVA22e
CsHVA22f
Seed-specific regulation
Elicitor-responsive element
Meristem expression
Auxin responsiveness
Anaerobic induction
Defense and stress responsiveness
Gibberellin-responsive
Zein metabolism regulation
Salicylic acid responsiveness
Low-temperature responsiveness
Drought-inducibility
Anoxic specific inducibility
MeJA-responsiveness
Light responsiveness
DNA binding protein
Abscisic acid responsiveness

Species	Motif (E-value)	Consensus sequence
C. clementina	4.0e-086	
	4.4e-059	
	1.4e-038	
	3.1e-008	
	4.1e-003	
C. sinensis	1.2e-130	
	1.7e-036	
	3.9e-032	
	4.2e-026	
	1.5e-019	
	2.1e-012	
	1.1e-006	
	3.3e-003	


CsHVA22d*
CsHVA22d
CcHVA22d
AtHVA22d
AtHVA22e
CcHVA22e
CsHVA22e
CcHVA22f
CsHVA22f
AtHVA22c
CsHVA22b
CsHVA22b*
CcHVA22b
AtHVA22a
CcHVA22c
CcHVA22c*
CsHVA22c
AtHVA22b
CcHVA22a
CsHVA22a
CsHVA22a*
33
100
100
96
40
100
63
100
61
24
33
100
38
86
99
32
100
34
0.9


Expression
0 5 10 15 20
CsHVA22a
CsHVA22c
CsHVA22d
CsHVA22b
CsHVA22e
Leaf
Flower
Callus
Fruit


A
WT
L1
L2
L3
L4
L5
B
Leaf disc dehydration rate (g)
Time (min)
WT
L1
L2
L3
L4
L5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
0
30
60
90
120
150
180
210
240
C
WT
L1
L2
L3
L4
L5