The selenium-binding protein of *Theobroma cacao*: an important protein to control biotic and abiotic stresses in plants a thermostable protein involved in the witches' broom disease resistance

Akyla Maria Martins Alves[1], Sara Pereira Menezes[1], Eline Matos Lima[1], Karina Peres Gramacho[2], Bruno Silva Andrade3, Monaliza Macêdo Ferreira[1], Carlos Priminho Pirovani[1], 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, 45662-900 Ilhéus-BA, Brazil.

[2]Cocoa Research Center, CEPLAC/CEPEC, 45600-970 Itabuna-BA, Brasil.

[3]Universidade Estadual do Sudoeste da Bahia (UESB), Av. José Moreira Sobrinho, Jequié, Bahia, 45206-190, Brazil

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

***Corresponding author:** Dr Fabienne Micheli, UESC, DCB, Rodovia Ilhéus-Itabuna km16, 45662-900, Ilhéus-BA, Brazil. Phone: +55 73 3680 5196. Fax: +55 73 3680 5226. E-mail: fabienne.micheli@cirad.fr

Abstract

The selenium-binding proteins are known to be inducers of apoptosis in human and animals, and have been studied as target for the treatment of various types of cancer. In plants, *SBP* expression has been related to abiotic and biotic stress resistance. The SBP from *Theobroma cacao* (*TcSBP*) was first identified from a cocoa-*Moniliophthora perniciosa* cDNA library. The present study provides details on the TcSBP gene and protein structure. Multiple alignments revealed conserved domains between SBP from plants, human and archea. Specific *cis*-elements related to hormone signaling and response to light were identified in the *TcSBP* promotor region. Homology modeling and molecular docking were performed and showed that the TcSBP has affinity to selenite in the active CSSC site. This result was confirmed by circular dichroism of the recombinant TcSBP, which also presented thermostable behavior. RT-qPCR analysis showed that *TcSBP* was differentially expressed in resistant *vs* susceptible cacao varieties inoculated by *M. perniciosa* and its expression was probably due to hormone induction via *cis*-regulating elements present in its promotor. The presence of the CSSC conserved domain in TcSBP structure suggested that *TcSBP* acted by altering oxidation/reduction of proteins during H_2O_2 production and programmed cell death (PCD) in the final stages of the witches' broom disease. Moreover, TcSBP was able to bind *in silico* the selenite and potentially may be able to bind heavy metals. For these reason, the TcSBP appeared to be a good candidate for biotechnological applications aiming control of biotic and abiotic stresses in plants. To our knowledge, this is the first *in silico* and *in vitro* analysis of the SBP from cacao.

Keywords: *Moniliophthora perniciosa*, selenium, molecular docking, gene expression

1. Introduction

Cacao is an endemic plant of tropical forests in South America and is grown mainly for the production of cocoa liquor, butter and powder for the chocolate industry [1]. However, cocoa production has been severely damaged by diseases caused by fungi and oomycetes [2]. In Brazil, the witches' broom disease caused by the basidiomycete *M. Moniliophthora perniciosa* caused drastic economic and social changes in the affected areas [2]. *Moniliophthora perniciosa* has two distinct phases of development: a biotrophic phase characterized by a monokaryotic and intercellular mycelium, and a necrotrophic phase characterized by a dikaryotic and intracellular mycelium [3, 4]. The transition from biotrophic to necrotrophic phase is characterized by significant accumulation of H_2O_2 and by programmed cell death (PCD). In cacao plants susceptible to *M. perniciosa*, the PCD occurs initially as a defense mechanism and then is deflected by the fungus for its own profit, allowing sporulation and further propagation [3]. Thus, the study of compounds, genes and proteins that induce PCD and/or facilitate facilitating the transition from biotrophic to necrotrophic phase [3, 4] is important to better understand the cacao-*M. perniciosa* pathosystem and to so develop strategies for witches' broom disease control.

The selenium-binding protein (SBP) is known as a tumor suppressor involved in the regulation of cell proliferation, senescence, migration and apoptosis [5] and has been studied as a target for the treatment of various types of cancer in animals and human, generally in related studies [6]. In human, it has been reported that the expression of the *SBP* gene was reduced in tumor tissues when compared to healthy ones, and for this reason, this gene has been indicated as a good predictor of clinical outcome of cancer [5]. SBP was also was related to resistance to abiotic and biotic stresses in plants [7, 8]. In several plant species, studies showed the correlation between SBP level and tolerance

to metal toxicity – e.g. by Se/selenite (SeO_3^{2-}), cadmium and sulfur – suggesting that SBP participated to metal detoxification mechanism. In several plant species, studies showed the correlation between SBP level and tolerance to metal toxicity (e.g. by Se/selenite SeO_3^{2-}, cadmium and sulfur) suggesting that SBP participated to metal detoxification mechanism [8]. In Arabidopsis, SBP1 was considered a general stress-responsive gene [9]. In transgenic rice plants (*Oryza sativa*), the overexpression of *OsSBP* increased the plant resistance to the rice blast disease caused by the fungus *Pyricularia grisea*, and to the bacterial blight due to *Xanthomonas oryzae* pv. oryzae [7, 10]. The SBP function was related to its ability to bind covalently Se covalently [5] and to the regulation of Se and SeO_3^{2-} metabolism [5].

The selenium-binding protein gene from *Theobroma cacao* (*TcSBP*) was first identified from a cocoa-*Moniliophthora perniciosa* cDNA library [11]. Here, we identified and analyzed *in silico* and *in vitro* the unique SBP sequence found in the *T. cacao* genome. The present study provides details on the *TcSBP* gene, including promotor analysis, as well as on the corresponding protein. The protein modeling and docking analyses revealed that TcSBP interacts with selenite at its CSSC conserved domain. Moreover, we showed that *TcSBP* presented differential expression pattern between resistant and susceptible cacao plants infected by *M. perniciosa*. The expression results associated to the structural analysis of the gene and protein, allowed a better understanding of the SBP role in cacao plants submitted to biotic and abiotic stresses. To the best of our knowledge, this is the first study of SBP in plants infected by hemibiotrophic fungi. The aim of this study was to understand the role of TcSBP during the cacao-*M. perniciosa* interaction as well as to identify possible action mechanisms of this protein associated to its ligation to selenite.

2. Methods

2.1 *In silico* analysis of TcSBP

The *TcSBP* cDNA was identified from a library of *T. cacao* meristem (genotype TSH1188) infected by *M. perniciosa* [11]. The complete sequence of the entire gene and of the corresponding proteins was obtained from the CocoaGenDB database (http://cocoa-genome-hub.southgreen.fr/). Open reading frame (ORF) detection was performed using the ORFinder software (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Sequence homology search was made with BLAST [12] on the National Center for Biotechnology Information (NCBI). Multiple sequence alignment was performed with the ClustalW2 software [13]. The prediction of theoretical isoelectric point (pI) and molecular weight (MW) was obtained using the Expasy Molecular Biology Server (www.expasy.org)ProtParam tool (https://web.expasy.org/protparam/ [14]). Post-translational events were predicted using the NetPhos 2.0 Server to identify putative sites of phosphorylation (Ser/Thr/Tyr) [15] and NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) to identify putative sites of N-glycosylation. Predictions of subcellular localization were conducted with the programs TargetP 1.1 server [16] and PSORT [17]. Signal peptide presence was analyzed using the SignalP 4.0 Server [18]. The TcSBP was were used, together with the sequences of the homolog proteins from other plants and mammals, to construct an unrooted phylogenetic tree by the neighbor joining method with the ClustalW2 [13] and MEGA 6 [19] programs. For the dendrogram analysis, only complete sequences with high identity (>70%) from plants and mammals were considered. The tree was constructed using the neighbor-joining method of ClustalW, with 1000 bootstraps. The search for regulatory sequences of *TcSBP* was performed in the CocoaGenDB database. A region of 1500 bp upstream the UTR was identified and selected for *cis*-element analysis. The presence of

126 the *cis*-regulatory elements in the promoter region of *TcSBP* was analyzed using the
127 plantCARE (sphinx.rug.ac.be:8080/PlantCARE/cgi/index.html) software [20].
128
129 **2.2 Molecular modeling of TcSBP and docking analysis**
130 The prediction of the three-dimensional (3-D) model of the TcSBP protein was
131 obtained using the Swiss-Model server and the Automated Protein Homology-
132 modeling, which relies on the high similarity of target-template [21]. The crystal
133 structure of the *Sulfolobus tokodaii* SBP was used as template (Protein Data Bank code:
134 2ECE) to build the structural model of TcSBP. The stereochemical quality of the
135 models was evaluated using the Procheck [22] and Anolea programs [23]. The 3-D
136 model visualization was obtained using the PyMol (The PyMOL Molecular Graphics
137 System, Version 1.5.0.4 Schrödinger, LLC.) and Discovery Studio 4.0. For docking
138 analysis of TcSBP 3-D model (receptor) with inorganic selenium compounds (ligands),
139 selenite and sodium selenite were search by keyword in the PubChem databases
140 (http://pubchem.ncbi.nlm.nih.gov). The 2D structure of these compounds were copied
141 in Similes and saved in PDB format using the Marvin program, and then used for
142 molecular docking. The ligand and receptor molecules were prepared using the
143 AutoDockTools 1.5.6 [24]. The grid definition, adjusted to active site, was set up
144 manually according to the recommendations of the program manual [24, 25]. The ligand
145 and receptor structures were then saved in pqbqt format to be used for docking
146 calculations. AutoDock Vina was used to perform Docking Scoring for each ligand-
147 receptor complex [25]. Before running each Docking calculation, a configuration file
148 was generated with information about grid size and coordinates, and indicating the
149 ligand and receptor files. The reports for each calculation were analyzed to obtain
150 affinity energy (kCal/Mol) values for each ligand conformation in its respective

151 complex. PyMol program was used to verify the ideal complex by considering all
152 stereochemical aspects previously evaluated and the free-energy results. Then the best
153 ligand i.e., that fit best in the active site, was selected.
154
155 **2.3 Plant material**
156 Plant material was obtained as previously described [26]. Seeds of *T. cacao*
157 genotypes Catongo (susceptible to *M. perniciosa*) and TSH1188 (resistant to *M.*
158 *perniciosa*) were germinated and grown at CEPLAC/CEPEC (Ilhéus, Bahia, Brazil)
159 greenhouses. Twenty to thirty days after germination, the apical meristems of the
160 plantlets were inoculated by the droplet method with a basidiospore suspension of *M.*
161 *perniciosa* (inoculum from isolate 4145 maintained in the CEPLAC/CEPEC
162 phytopathological *M. perniciosa* collection under number 921 of the WFCC;
163 http://www.wfcc.info/index.php/collections/display). After inoculation, the plantlets
164 were kept for 24 h at 25 ± 2°C and 100% humidity. Rate of disease fixation based on
165 presence/absence of symptoms in each genotype, was evaluated 60 days after
166 inoculation (dai); disease rate was 45% and 80% for TSH1188 and Catongo,
167 respectively. Moreover, the presence of *M. perniciosa* in the plant material was checked
168 by semi-quantitative RT-PCR using specific *M. perniciosa* actin primers [26]; both
169 genotypes presented fungus incidence (data not shown) coherent with previous data
170 obtained in the same conditions of plant culture and inoculation [4, 26]. Apical
171 meristems were harvested at 24, 48 and 72 hours after inoculation (hai), and 8, 15, 30,
172 45, 60 and 90 dai. Non-inoculated plants (controls) were kept and harvested under the
173 same conditions at 72 hai, 60 dai and 90 dai. For each genotype and at each harvesting
174 time (for inoculated and non-inoculated plants), 20 samples were collected (1 sample =
175 1 apical meristem of 1 cacao plantlet). The 20 samples collected from one genotype at

176 one harvesting time were pooled; thus 9 inoculated and 5 non inoculated (control)
177 samples were immediately frozen in liquid nitrogen and stored at -80°C until use.
178 Pooling samples before RNA extraction has the advantage of reducing the variation
179 caused by biological replication and sample handling [26].
180
181 **2.4 Reverse transcription quantitative PCR analysis**
182 Total RNA was extracted from macerated samples using the RNAqueous Kit®
183 (Ambion) according to the manufacturer's instructions, with modifications as
184 previously described [26]. The synthesis of cDNA was carried out using Revertaid First
185 Fisrt Strand cDNA Synthesis Kit according to the manufacturer's instructions (Thermo
186 Scientific). The cDNA quantification was carried out on the GeneQuant pro UV/Vis
187 spectrophotometer (Amersham). For the qPCR analysis, two cacao endogenous
188 reference genes were used: the malate dehydrogenase (MDH) and β-actin (ACT),
189 previously identified as *T. cacao* housekeeping genes and tested in our experimental
190 conditions (same plant material, same equipment [26]; Supplementary material 1).
191 Specific *TcSBP* primers were also defined (Supplementary material 1). The expression
192 analysis of *TcSBP* was performed using standard settings of the ABI PRISM 7500 and
193 Sequence Detection System (SDS) software, version 1.6.3 (Applied Biosystems). The
194 qPCR reaction consisted of 10 ng/µl of cDNA, 1 µM of each primer from reference or
195 targets genes (Supplementary material 1) and 11 µl of Power SYBR Green Master Mix
196 (Applied Biosystems) in a total volume of 22 µl. Cycling conditions were: 50°C for 2
197 min then 95°C for 10 min, followed by 40 cycles at 95°C for 15 s, 60°C for 30s and
198 60°C for 1 min. To verify that each primer pair produced only a single PCR product, a
199 dissociation analysis was carried out from 60°C to 95°C and analyzed with the
200 Dissociation Curve 1.0 program (Applied Biosystems). The expression analysis of

TcSBP in meristems of cacao plantlets inoculated or not with *M. perniciosa*, was obtained by RT-qPCR using 6 experimental technical replicates. The relative expression was analyzed with the comparative Ct method ($2^{-\Delta\Delta Ct}$) using MDH and ACT as endogenous reference genes and non-inoculated plants (control) as calibrator. Statistical analysis was made using the SASM-Agri software which tested the experiments as a completely randomized design. *t*-test and *F*-test (ANOVA) were applied with a critical value of 0.01. The Scott-Knott ($P \leq 0.01$) test was employed for mean separation when *F*-values were significant.

2.5 Obtaining of the recombinant TcSBP

Total RNA was extracted as described above (§ 2.4) from a pool of all samples of TSH1188 *T. cacao* seedlings inoculated with the fungus *M. perniciosa* (see §2.3 and *TcSBP* expression data). The synthesis of cDNA was carried out as described above (§ 2.4). To develop a *TcSBP* cloning strategy in pET-28a(+) vector (used for heterologous expression), specific primers (F 5'-AAGACAGGGCATATGGCCGGTAACG-3' and R 5'-CGGTAAAGAGTGAAGCTTCACAGCAAG-3' containing *Nde*I and *Hind*III restriction sites) were designed using the Custom Primers-OligoPerfect ™ Designer (Invitrogen), OligoAnalyser 3.1 (IDT SciTools/Integrated DNA Technologies, Inc., California), and Webcutter (http://heimanlab.com/cut2.html) programs. The *TcSBP* amplification reaction was performed in a total volume of 25 µl containing 1 µl of the cDNA sample, 2.5 µl of 10X PCR buffer, 0.75 µl of $MgCl_2$ (1.5 mM), 1 µM of each primer (0.2 mol/µl), 1U Recombinant Taq DNA Polymerase (Invitrogen) and ultrapure water (qsp 25 µl). The reaction was conducted in a Veriti 96-well Thermal Cycler (Applied Biosystems) and consisted of the following steps: 4 min at 94°C; 30 seconds at 94°C; 1 min at 58ºC; 4 min at 72°C; 7 min at 72°C; final wait of 15ºC. The PCR

product was analyzed on a 1% agarose gel stained with GelRed (Biotium). Purification of the PCR product was performed using the GFX™ PCR DNA and Gel Band Purification Kit according to the manufacturer's recommendations (GE Healthcare). The purified PCR product, as well as the pET-28a(+) vector, were digested at 37°C for 4 h by *Nde*I and *Hind*III in a final volume reaction of 120 µL containing 96 µl of DNA, 12 µl of the Tango buffer, 4.5 µL of the each enzyme and 3 µl of ultrapure water. After purification using the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare), the reactions were treated with the FastAP enzyme according to the manufacturer's recommendations (Thermo Fisher Scientific). Cloning step was performed using 90 µl of insert, 10 µl of pET-28a(+) vector, 10 µl of T4-DNA ligase, 15 µl of 10X buffer and 25 µl of ultrapure water. Cloning reaction was used to transform *Escherichia coli* C43 strain by heat shock method [27] and the transformed colonies were selected on LB medium containing kanamycin (50 µg/ml). Transformed colonies were confirmed by PCR using the same primers and conditions as described above. For protein induction, PCR-confirmed colony was grown on LB medium at 37°C under agitation at 180 rpm until OD_{600nm} was reached 0.6 to 1.0, then incubated with 0.4 µmol/l of IPTG at 18°C for 16 h under agitation, and finally centrifuged at 180 rpm. The pellet was ressuspended in 50 µg/ml of lysozyme, kept on ice for 20 min, and then sonicated in an ice bath with 10 pulses of 10s (on), 25 pulses of 10 seconds (off) and 70% amplitude, until breaks the pellet viscosity. The lysate material was centrifuged at 5000 rpm, 4°C for 45 min and the supernatant was used for recombinant protein purification on TALON Superflow cobalt ion (Co^{2+}) (GE Healthcare). The recombinant protein was eluted with 150 mM of imidazole and submitted to dialysis in phosphate buffer (5mM). rTcSBP purification was checked on 15% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as previously described [28].

Electrophoresis ran for 2 h at 150 V with the cube immersed in ice, and the proteins were detected on electrophoresis gel by Coomassie Blue G250 staining [29]. The concentration of the rTcSBP was determined by the Bradford method [30] on the microplate reader ELX 800 (Biotek) using the bovine serum albumin (BSA) for the standard curve calculation.

Circular dichroism

The circular dichroism (CD) was carried out on a spectropolarimeter (model J-815, Jasco, MD) equipped with a Peltier temperature controller and a cell support with a thermostat connected to a thermostatic bath. The CD of rTcSBP was performed at a concentration of 0.355 μg μL^{-1} of the protein in sodium phosphate buffer (10 mmol L^{-1}, pH 7.0). Three CD spectra (experimental replicates) of rTcSBP were measured in wavelength range of 190-240 nm at 25°C and 95°C. The thermal stability analysis of rTcSBP was also evaluated using unfolding (refraction denaturation from 25 to 95ºC) and refolding (variation from 95 to 25ºC) at 205 nm in sodium phosphate buffer (10 mmol L^{-1}, pH 7.0). The rTcSBP structure was also analyzed in presence of sodium selenite (1:1 molar ratio) at 190-240 nm and 25°C. The secondary α-helix and β-sheet structures were estimated using the K2D2 server (http://cbdm-01.zdv.uni-mainz.de/~andrade/ k2d2/).

3. Results

3.1 *In silico* analysis of *TcSBP*

In silico analysis on CocoaGenDB revealed the presence of a unique sequence encoding TcSBP located on the chromosome 4 (Fig.1). The *TcSBP* gene is 4774 bp in length, contains 7 exons and 6 introns and has an ORF of 1431 bp (Fig.1). *TcSBP*

276 encodes a protein of 476 amino acids, with a MW of 52732.7 Da and a putative pI of
277 5.74. Twenty-one putative phosphorylation and no N-glycosylation sites were observed
278 in the TcSBP protein sequence (Fig. 2A). TcSBP does not contain any signal peptide,
279 and its subcellular location is unknown (data not shown). The alignment of TcSBP with
280 SBPs from other organisms revealed high identity (>70%) with plant species and human
281 and allowed the identification of the CSSC, DELHH and HGD conserved region
282 (Fig.2A). The full-length amino acid sequences of plants and mammalian SBPs
283 available in the databases were used to build an unrooted dendrogram (Fig. 3;
284 Supplementary material 2). Plant and mammalian SBPs were separated and formed two
285 distinct branches in the dendrogram (Fig. 23). *In silico* analysis of the *TcSBP* promoter
286 region revealed the presence of several conserved *cis*-elements characteristic of
287 different plant species (Table 1). Among these *cis*-elements, TATA-box and GA-motif
288 related with transcription start, and with *cis*-acting in promoter and enhancer regions,
289 respectively, were found (Table 1). Elements related to light response and circadian
290 control were encountered (CAAT-box and Circadian, respectively) as well as elements
291 related to response to abiotic stress or external stimuli (anaerobic induction/ARE, low-
292 temperature response/LTR) were found. Finally, several *cis*-elements related to
293 hormone signaling were found such as ABRE (abscisic acid responsiveness), TATC-
294 box (gibberellin responsiveness), TCA (salicylic acid responsiveness) and TGA-box
295 (auxin-responsive element) (Table 1).
296
297 **3.2 Homology modeling of TcSBP and molecular docking of TcSBP with selenite**
298 The modeling of the TcSBP was based in the X-ray structure of the SBP from
299 *Sulfolobus tokodaii* (StSBP). The alignment of the amino acid sequence of TcSBP with
300 the StSBP (Fig. 32A) presented 38% of identity, which is superior to the reliable cut-off

of 25% [31]. And the structure identity between StSBP and TcSBP was 43.91% (X-ray, 2.0Å; data not shown). The structural modeling of TcSBP revealed the presence of 7 α-helices and numerous β-strands and P-loops (Fig. 32B). The Ramachandran plots showed that 99.9% of the amino acids of the TcSBP model were located in the most favored regions (Supplementary material 3). The TcSBP showed important amino acid residues organized on domains (DELHH, HGD, CSSC/CSSA) based on its high identity (>70%) with SBPs from other species (Fig. 32A). The tridimensional structure of TcSBP obtained by homology modeling showed that the protein contains 8 α-helices (in red, Fig. 32B) and 30 β-strands (in blue, Fig. 32B). The CSSC conserved motif of TcSBP was showed in the surface format structure (in pink, Fig. 32C) and as stick representation (in pink, Fig. 32D). Putative phosphorylation sites were present at the superficies of the protein structure; interestingly, phosphorylation sites were present in the CSSC domain of the protein (Fig. 2C). The interaction between the conserved motif CSSC of TcSBP and the selenite occurred with a setting of -1.9 kCal/mol (Fig. 2E). No interaction between TcSBP and selenate was observed (data not shown).

3.3 Production of rTcSBP and circular dichroism

The rTcSBP was expressed in *E. coli* C43, purified on affinity column, and visualized on SDS-PAGE gel (see Material and Methods). The purified rTcSBP showed molecular weight (56 kDa) coherent with the predicted one (Supplementary material 4). Circular dichroism spectra of rTcSBP in phosphate buffer pH 7.0 showed that the protein contained both α-helix and β-strand elements (Fig. 4A). The rTcSBP structure spectra showed a lower peak at ~205 nm, mainly at 95°C (Fig. 4A). The rTcSBP secondary structure variation according to the temperature (unfolding and refolding) at 205 nm showed that the rTcSBP was duly folded at 25°C and denatured from 25°C to

95°C (unfolding, Fig. 4B), while the refolding occurred from 95°C to 35°C with a fully recovery of the secondary structure at 35°C (0% of denaturation; Fig. 4B). Circular dichroism of rTcSBP in presence and absence of sodium selenite showed that the sTcSBP interacts with selenite but that this interaction did not influence the protein structure (Fig. 4C).

3.34 Differential expression of *TcSBP* in resistant and susceptible cacao genotypes infected by *M. perniciosa*

The expression of the *TcSBP* gene was analyzed for each cacao genotype, i.e. TSH1188 (resistant to witches' broom disease) and Catongo (susceptible), infected or not (control) with *M. perniciosa* (Fig. 5). For both genotypes and for all the harvesting point, the PCR amplification occurred at the same and unique melting temperature for each gene showing that only the corresponding gene was amplified (Supplementary material 5). At 24 hai, *TcSBP* was significantly more expressed in inoculated susceptible variety than in the control, while its expression did not differ between inoculated resistant variety and the control plants. At 48 and 72 hai no significant difference of *TcSBP* expression between the resistant and susceptible varieties was observed, however, at 72 hai, *TcSBP* had a lower expression in the inoculated plants (whatever the genotype) than in the control plants. At 8 dai, the resistant variety showed the highest expression of *TcSBP* in relation to the other times after inoculation. At this time the susceptible variety exhibits a relative *TcSBP* expression lower than in the control plants. At 15 dai, the relative *TcSBP* expression in the resistant variety did not differ from the control, while in the susceptible variety the expression was lower than in the control plants. At 30 and 45 dai, the *TcSBP* expression was also significantly different between the varieties; *TcSBP* presented significantly higher expression in the

susceptible variety. At 30 and 45 dai, in the resistant variety, *TcSBP* expression did not differ between inoculated and the control plants. At 60 and 90 dai, no significant difference in *TcSBP* expression was observed in the studied varieties.

4. Discussion

Here, we characterized *in silico* a *SBP* gene from *T. cacao* (Fig. 1 and 2). This gene is present in the cacao genome in a unique copy located in the chromosome 4 (Fig. 1). The corresponding protein (476 amino acids; Fig. 2A) presented a secondary structure containing α-helices and β-strands (Fig. 2B and Fig4A) and is thermostable *in vitro* (Fig. 2A and B). Interestingly, TcSBP showed a high level of identity with the SBP from *Sulfolobus tokodaii* (Fig. 2A), which is a thermophilic archaea able to survive in high temperature environments (e.g. 70°C) [32]. The promotor region of *TcSBP* contained several *cis*-elements related to biotic stress and to hormone signaling such abscisic acid, gibberellin, salicylic acid and auxin (Table 1). Some of these hormones are known to be highly involved in witches' broom disease development [11, 33-35]. Several genes related to plant hormone signaling were altered in cacao plants in response to *M. perniciosa* infection, such as gene involved in gibberellin response as well as auxin-responsive genes, both up-regulated in green broom [35]. Is has been suggested that the up-regulation of auxin-responsive genes in infected cacao plants is a consequence of the biosynthesis of this hormone by *M. perniciosa* [33-35]. In various plant pathogens, the production of auxin has been correlated to the reduction of the plant defenses mediated by salicylic acid, and consequently to the pathogenicity and to the biotrophic stage development [33, 35]. Here, we observed that the *TcSBP* gene was significantly more expressed in the susceptible variety during both the asymptomatic and the green broom stages (Fig. 4). Thus, the *TcSBP* expression may be related to its

induction by hormones present at these stages as the consequence of the infection by the fungus (Fig. 4), as previously observed for SBP genes from other species [10]. In rice, the SBP expression increased rapidly in response to jasmonic acid and salicylic acid and also responded to abscisic acid and paraquat, both responsible for ROS reactive oxygen species generation [10].

The witches' broom disease development has been previously associated to increased H_2O_2 levels in both resistant and susceptible cacao varieties, however with distinct temporal and functional patterns [3, 36]. Once produced at the beginning of the infection by the resistant variety, H_2O_2 contributes to the infection control and to plant resistance. In advanced stages of the disease in the susceptible variety, it promotes the pathogen development and the finalization of its life cycle [36]. Interestingly, the *TcSBP* expression also presented distinct temporal patterns; it was higher at 8 and 15 dai in the resistant variety, and from 30 dai until the end of the disease in the susceptible variety (Fig. 4). Considering that TcSBP contains a CSSC conserved domain (*in silico* analysis, Fig. 2E), which is known to acts as an active redox center controlling oxidation/reduction of protein *in vivo* [10], it could be suggested that TcSBP gene may play a role in protecting cells from oxidative stress in swelling ramification phase in the resistant genotype, and regulating the PCD in the susceptible one during the green broom stage (Fig. 4; [3, 36]). This results could be related to the known SBP function, i.e. its capacity to inhibit the cell proliferation, and consequently its involvement in apoptosis mechanism [5]. In rice, *OsSBP* was involved in the accumulation of reactive oxygen intermediates produced by NADPH oxidase, and the observed H_2O_2 accumulation was due to the reduction of scavenging enzyme activity [7].

Finally, the TcSBP protein contains multiple predicted phosphorylation sites (Fig. 2A); protein phosphorylation is a ubiquitous mechanism for the temporal and

spatial regulation of proteins involved in almost every cellular process [37]. *TcSBP* contains three domains, DELHH, HGD and CSSC, highly conserved among the SBPs from other organisms (Fig. 2A); the CSSC domain contained predicted phosphorylation sites (Fig. 2C). It has been suggested that the characteristic CSSC motif has a strong affinity for selenium and heavy metals [38]. In *T. cacao*, we observed that the conserved domain CSSC was able to bind, *in silico*, the selenite (Fig. 2E). The binding of rTcSBP with the selenite was confirmed *in vitro* by circular dichroism, showing that this interaction did not modify the protein structure (Fig. 4C). Some phosphorylation sites (the stable ones) are known to be important for the regulatory function of the protein [39] and, here, it may be suggested that the phosphorylation at the CSSC domain may be related to the binding with selenite, as well as to the protein function as regulator of oxidation/reduction during the plant-pathogen interaction. Moreover, the conserved domain CSSC potentially may be a binding site for other metals as observed in other organisms [38]. This could be interesting considering the concerns about high concentrations of heavy metals (e.g. Pb, Cd, Cu) found in beans of cacao as well as in its products (including chocolate). Contamination of cacao is attributed to the increasing use of phosphate fertilizers but also to several kinds of industrial processes close to the cacao production areas [40].

5. Conclusion

Here we showed the first *in silico* and *in vitro* analysis of the SBP from cacao. The rTcSBP is a thermostable protein able to bind *in silico* and *in vitro* the selenite and potentially may be able to bind heavy metals. *TcSBP* is differentially expressed in resistant *vs* susceptible cacao varieties inoculated by *M. perniciosa* and its expression is probably due to hormone induction (e.g. IAAindole acetic acid, ASsalicylic acid,

426 gibberellins) via *cis*-regulating elements present in its promotor (e.g. TATC-box, TCA,
427 TGA-box). The presence of the CSSC conserved domain in TcSBP structure suggested
428 that *TcSBP* acted by altering oxidation/reduction of proteins during H_2O_2 production
429 and PCD in the final stages of the witches' broom disease. Moreover, TcSBP was able
430 to bind *in silico* the selenite and potentially may be able to bind heavy metals. For these
431 reasons, the TcSBP appeared to be a good candidate for biotechnological applications
432 aiming control of biotic (e.g. pathogens) and abiotic (e.g. heavy metals) stresses.
433
434 **Availability of supporting data**
435 The data sets supporting the results of this article are included within the article and its
436 additional files.
437
438 **Abbreviations**
439 ACT: actin; dai: days after inoculation; hai: hour after inoculation; pI: isoeletric point;
440 MDH: malate dehydrogenase; MW: molecular weight; ORF: open reading frame; PCD:
441 programmed cell death; SBP: selenium-binding protein; Se: selenium; UTR:
442 untranslated regions.
443
444 **Competing interest**
445 No conflicts of interest to declare.
446
447 **Authors' contribution**
448 AMMA was responsible for the execution of all the experimental steps; AMMA, SPM,
449 BSA and FM analyzed and discussed the data; AMMA and FM wrote the manuscript;
450 SPM and EML gave support in qPCR experiment; BSA gave support in modeling and

451 docking analyses; KPG was responsible for plant material production and inoculation
452 with *M. perniciosa*; MMF and CPP supported the circular dichroism analysis; FM and
453 BSA were responsible for the conception and design of the experiments; FM was
454 responsible for the financial support of the research and for the advising of AMMA,
455 EML and SPM.
456
457 **Acknowledgements**
458 The work of AMMA was supported by the Fundação de Amparo à Pesquisa do Estado
459 da Bahia (FAPESB). The work of SPM, EML and MMF was supported by the
460 Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). FM, KPG
461 and CPP were supported by a Productivity grant from Conselho Nacional de
462 Desenvolvimento Científico e Tecnológico (CNPq). This research was supported by the
463 FAPESB project (DTE0038/2013) coordinated by FM. The authors thank Francisca
464 Feitosa Jucá (UESC/Ceplac) and Louise Araújo Sousa (Ceplac) for technical help in
465 plant inoculation experiments, Dr. Gesilvado Santos (UESB) for advices and discussion
466 about SBP role. This research was made in the frame of the International Consortium in
467 Advanced Biology (CIBA).
468
469 **References**

470 [1] G.A.R. Wood, R.A. Lass, Cocoa, Fourth Edition ed., Blackwell Science Ltd., 1985.
471 [2] H.C. Evans, Cacao Diseases—The Trilogy Revisited, Phytopathology, 97 (2007) 1640-1643.
472 [3] G.O. Ceita, J. Macedo, T. Santos, L. Alemanno, A.S. Gesteira, F. Micheli, A. Mariano, K.
473 Gramacho, D.C. Silva, L. Meinhardt, Involvement of calcium oxalate degradation during
474 programmed cell death in *Theobroma cacao* tissues triggered by the hemibiotrophic fungus
475 *Moniliophthora perniciosa*, Plant Science, 173 (2007) 106-117.
476 [4] K. Sena, L. Alemanno, K.P. Gramacho, The infection process of *Moniliophthora perniciosa* in
477 cacao, Plant Pathology, (2014) 1272–1281.
478 [5] S. Zhang, F. Li, M. Younes, H. Liu, C. Chen, Q. Yao, Reduced Selenium-Binding Protein 1 in
479 Breast Cancer Correlates with Poor Survival and Resistance to the Anti-Proliferative Effects of
480 Selenium, PLoS ONE, 8 (2013) e63702.

[6] E. Ansong, W. Yang, A.M. Diamond, Molecular cross-talk between members of distinct families of selenium containing proteins, Molecular nutrition & food research, 58 (2014) 117-123.
[7] K. Sawada, M. Hasegawa, L. Tokuda, J. Kameyama, O. Kodama, T. Kohchi, K. Yoshida, A. Shinmyo, Enhanced Resistance to Blast Fungus and Bacterial Blight in Transgenic Rice Constitutively Expressing OsSBP, a Rice Homologue of Mammalian Selenium-binding Proteins, Bioscience, Biotechnology, and Biochemistry, 68 (2004) 873-880.
[8] A. Agalou, A. Roussis, H.P. Spaink, The Arabidopsis selenium-binding protein confers tolerance to toxic levels of selenium, Functional Plant Biology, 32 (2005) 881-890.
[9] C. Valassakis, P. Livanos, M. Minopetrou, K. Haralampidis, A. Roussis, Promoter analysis and functional implications of the selenium binding protein (SBP) gene family in Arabidopsis thaliana, Journal of Plant Physiology, 224-225 (2018) 19-29.
[10] K. Sawada, L. Tokuda, A. Shinmyo, Characterization of the Rice Blast Fungal Elicitor-Responsive Gene OsSBP Encoding a Homolog to the Mammalian Selenium-binding Proteins, Plant Biotechnology, 20 (2003) 177-181.
[11] A.S. Gesteira, F. Micheli, N. Carels, A. Da Silva, K. Gramacho, I. Schuster, J. Macedo, G. Pereira, J. Cascardo, Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa*, Ann Bot, 100 (2007) 129-140.
[12] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res, 25 (1997) 3389-3402.
[13] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, Bioinformatics, 23 (2007) 2947-2948.
[14] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, B. A., Protein Identification and Analysis Tools on the ExPASy Server, in: J.M.W. (ed) (Ed.) The Proteomics Protocols Handbook, Humana Press 2005, pp. 571-607.
[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] O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools, Nat. Protocols, 2 (2007) 953-971.
[17] K. Nakai, P. Horton, PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization, Trends in Biochemical Sciences, 24 (1999) 34-35.
[18] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat Meth, 8 (2011) 785-786.
[19] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0, Molecular Biology and Evolution, 30 (2013) 2725-2729.
[20] 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.
[21] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics, 22 (2006) 195-201.
[22] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, Journal of Applied Crystallography, 26 (1993) 283-291.
[23] F. Melo, E. Feytmans, Assessing protein structures with a non-local atomic interaction energy, Journal of Molecular Biology, 277 (1998) 1141-1152.
[24] M. Sanner, A.J. Olson, J.C. Spehner, Reduced Surface: an Efficient Way to Compute Molecular Surfaces, Biopolymers, 38 (1996) 305-320.
[25] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, Journal of Computational Chemistry, 31 (2010) 455-461.

[26] S. Pereira Menezes, E. de Andrade Silva, E. Matos Lima, A. Oliveira de Sousa, B. Silva Andrade, L. Santos Lima Lemos, K. Peres Gramacho, A. da Silva Gesteira, C. Pirovani, F. Micheli, The pathogenesis-related protein PR-4b from *Theobroma cacao* presents RNase activity, Ca^{2+} and Mg^{2+} dependent-DNase activity and antifungal action on *Moniliophthora perniciosa*, BMC Plant Biol, 14 (2014) 1-17.
[27] J. SAMBROOK, E.F. FRITSCH, E.F.T. MANIATIS, Molecular cloning - a laboratory manual, Cold Spring Harbor, New York, 1989.
[28] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 227 (1970) 680-685.
[29] V. Neuhoff, N. Arold, D. Taube, W. Ehrhardt, Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250., Electrophoresis, 9 (1988) 255-262.
[30] N.J. Kruger, The Bradford Method for Protein Quantitation, in: J.M. Walker (Ed.) Basic Protein and Peptide Protocols, Humana Press, Totowa, NJ, 1994, pp. 9-15.
[31] T. Schwede, J.r. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, Nucleic Acids Research, 31 (2003) 3381-3385.
[32] N. Ohtani, H. Yanagawa, M. Tomita, M. Itaya, Cleavage of double-stranded RNA by RNase HI from a thermoacidophilic archaeon, Sulfolobus tokodaii 7, Nucleic Acids Research, 32 (2004) 5809-5819.
[33] A. Kilaru, B.A. Bailey, K.H. Hasenstein, *Moniliophthora perniciosa* produces hormones and alters endogenous auxin and salicylic acid in infected cocoa leaves, FEMS Microbiology Letters, 274 (2007) 238-244.
[34] F.C. Chaves, T.J. Gianfagna, Necrotrophic phase of Moniliophthora perniciosa causes salicylic acid accumulation in infected stems of cacao, Physiological and Molecular Plant Pathology, 69 (2006) 104-108.
[35] P.J.P.L. Teixeira, D.P.d.T. Thomazella, O. Reis, P.F.V. do Prado, M.C.S. do Rio, G.L. Fiorin, J. José, G.G.L. Costa, V.A. Negri, J.M.C. Mondego, P. Mieczkowski, G.A.G. Pereira, High-resolution transcript profiling of the atypical biotrophic interaction between *Theobroma cacao* and the fungal pathogen *Moniliophthora perniciosa*, The Plant Cell, 26 (2014) 4245-4269.
[36] C.V. Dias, J.S. Mendes, A.C. dos Santos, C.P. Pirovani, A. da Silva Gesteira, F. Micheli, K.P. Gramacho, J. Hammerstone, P. Mazzafera, J.C. de Mattos Cascardo, Hydrogen peroxide formation in cacao tissues infected by the hemibiotrophic fungus *Moniliophthora perniciosa*, Plant Physiol Biochem, 49 (2011) 917-922.
[37] Samuel M. Pearlman, Z. Serber, James E. Ferrell, Jr., A Mechanism for the Evolution of Phosphorylation Sites, Cell, 147 (2011) 934-946.
[38] C.-L. Wu, W.-B. Zhang, K.-S. Mai, X.-F. Liang, W. Xu, J. Wang, H.-M. Ma, Molecular cloning, characterization and mRNA expression of selenium-binding protein in abalone (Haliotis discus hannai Ino): Response to dietary selenium, iron and zinc, Fish & Shellfish Immunology, 29 (2010) 117-125.
[39] F. Ardito, M. Giuliani, D. Perrone, G. Troiano, L. Lo Muzio, The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review), Int J Mol Med, 40 (2017) 271-280.
[40] E. Arévalo-Gardini, C.O. Arévalo-Hernández, V.C. Baligar, Z.L. He, Heavy metal accumulation in leaves and beans of cacao (Theobroma cacao L.) in major cacao growing regions in Peru, Science of The Total Environment, 605-606 (2017) 792-800.

Table 1. *Cis*-elements present in the promoter region of *TcSBP* gene.

***cis*-elements name**	**Organism**	**Sequence**	**Function**
ABRE	*Arabidopsis thaliana*	ACGTG	Abscisic acid responsiveness
	Hordeum vulgare	GCAACGTGTC	
	Arabidopsis thaliana	CACGTG	
ARE	*Zea mays*	AAACCA	Anaerobic induction
TATC-box	*Oryza sativa*	TATCCCA	Gibberellin-responsiveness
TCA-element	*Nicotiana tabacum*	CCATCTTTTT	Salicylic acid responsiveness
TGA-box	*Glycine max*	TGACGTAA	Auxin-responsive element
Circadian	*Lycopersicon esculentum*	CAAAGATATC	circadian control
CAAT-box	*Brassica rapa*	CAAAT	Light responsive
	Hordeum vulgare	CAAT	
	Glycine max	CAATT	
GA-motif	*Glycine max*	AAGGAAGA	*Cis*-acting in promoter and enhancer regions
LTR	*Hordeum vulgare*	CCGAAA	Low-temperature responsiveness
TATA-box	*Arabidopsis thaliana*	TATAAA	Transcription start
	Arabidopsis thaliana	TATAAA	
	Oryza sativa	TATAAGAA	
	Lycopersicon esculentum	TTTTA	
	Lycopersicon esculentum	TTTTA	
	Arabidopsis thaliana	TATAA	

Figure legends

Figure 1. Structural features of TcSBP sequence. Lengths of the regions are indicated at the scale.

Figure 2. Homology modeling of TcSBP and molecular docking. **A.** Alignment of the amino acid sequence of TcSBP (XP_007034202.1) with SBPs from *Homo sapiens* (HsSBP1, AAB02395.1, UniProtKB:Q13228) and *Sulfolobus tokodaii* (StSBP, WP_010977998.1, UniProtKB:Q976Y0). Gaps introduced to get the best alignment are indicated by (-), (*) represents identical amino acids between all sequences, (.) and (:) represent conserved substitutions and semi-conserved substitutions, respectively. The

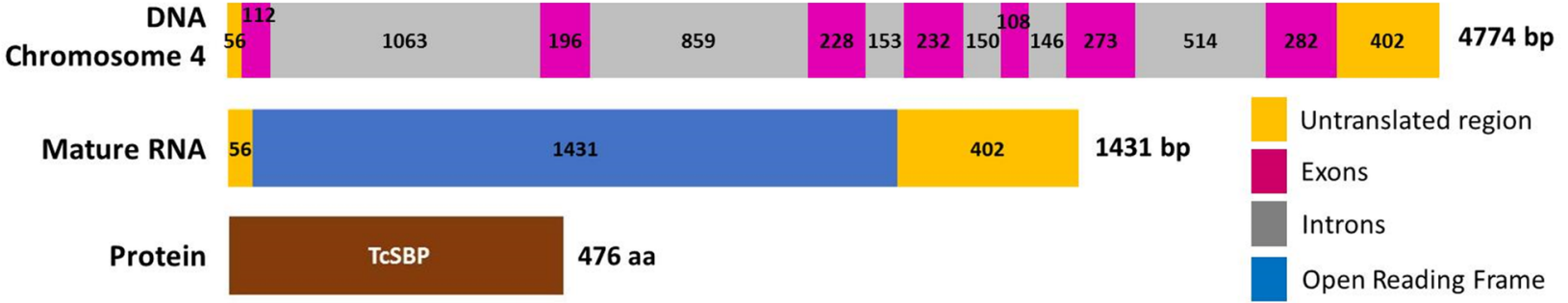
conserved motifs DELHH, CSSC and HGD are highlighted in blue, yellow and black, respectively. Phosphorylation sites are indicated by green squares. **B.** Tridimensional structure of TcSBP obtained by homology modeling with StSBP from *Sulfolobus tokodaii* (PDB code 2ECE.pdb). **C.** Surface representation of TcSBP with the indication of the conserved motifs CSSC in pink and of putative phosphorylation sites in green. **D.** Stereo view using sticks representation with focus on the conserved CSSS motif (in pink). **E.** Docking between TcSBP and selenite at the TcSBP CSSC site.

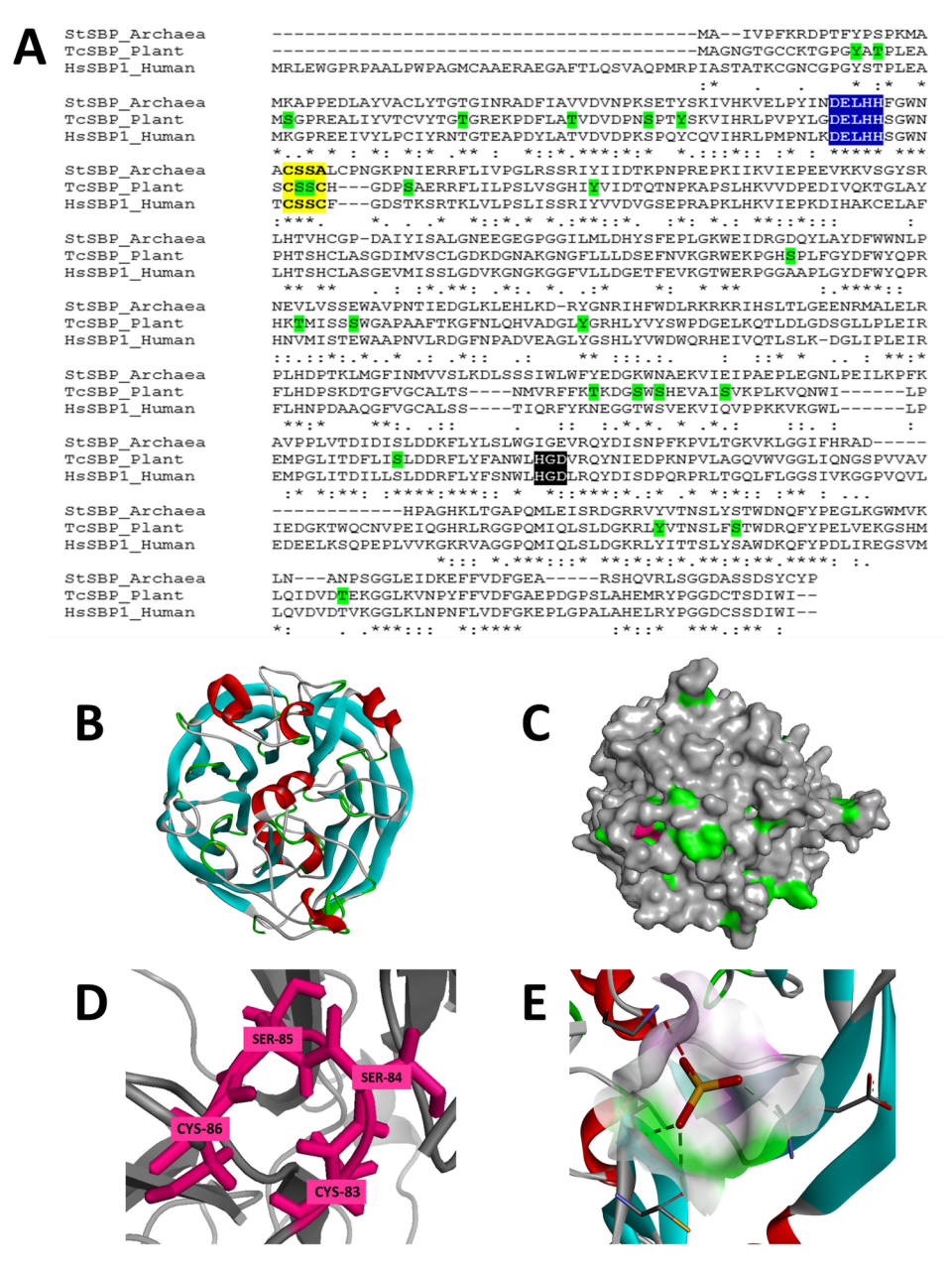
Figure 3. Dendrogram of TcSBP and SBP from other species. The green and red boxes indicate plant and mammalian SBPs, respectively. The amino acids sequences were derived from the GenBank under the following accession numbers: VvSBP, *Vitis vinífera* (XP_002267004.1; GI:225433514); CsSBP, *Citrus sinensis* (XP_006492530.1; GI:568879129); StSBP, *Solanum tuberosum* (XP_006360921.1; GI:565390388); PtSBP, *Populus trichocarpa* (XP_006373096.1; GI:566212227); RcSBP, *Ricinus communis* (XP_002520613.1; GI:255559184); TcSBP, *Theobroma cacao* (XP_007034202.1; GI:1063488813); CsSBP, *Cucumis sativus* (XP_004138463.1; GI:449441386); MsSBP, *Medicago sativa* (AJ401228.1; CAC67501.1); GmSBP, *Glycine max* (KHN24031.1; GI:734383687); PvSBP, *Phaseolus vulgaris* (XP_007163807.1; GI:593801540); ShSBP, *Sarcophilus harrisii* (XP_003773388.1XP_003769976.1; GI: 395535935); MmSBP, *Mus musculus* (AAA40104.1; GI:200952); RnSBP, *Rattus norvegicus* (NP_543168.1NP_001316822.1; GI: 1049480106); HgSBP, *Heterocephalus glaber* (XP_004875930.1XP_004854049.1; GI: 512984154); LaSBP, *Loxodonta africana* (XP_010587740.1; GI:731468243); BmSBP, *Bos mutus* (XP_005895013.1; GI:555965912); TtSBP, *Tursiops truncatus* (XP_004328520.1XP_019775099.1; GI:1131274601); PtSBP, *Pan troglodytes* (XP_001172033.2; GI:332810264); HsSBP,

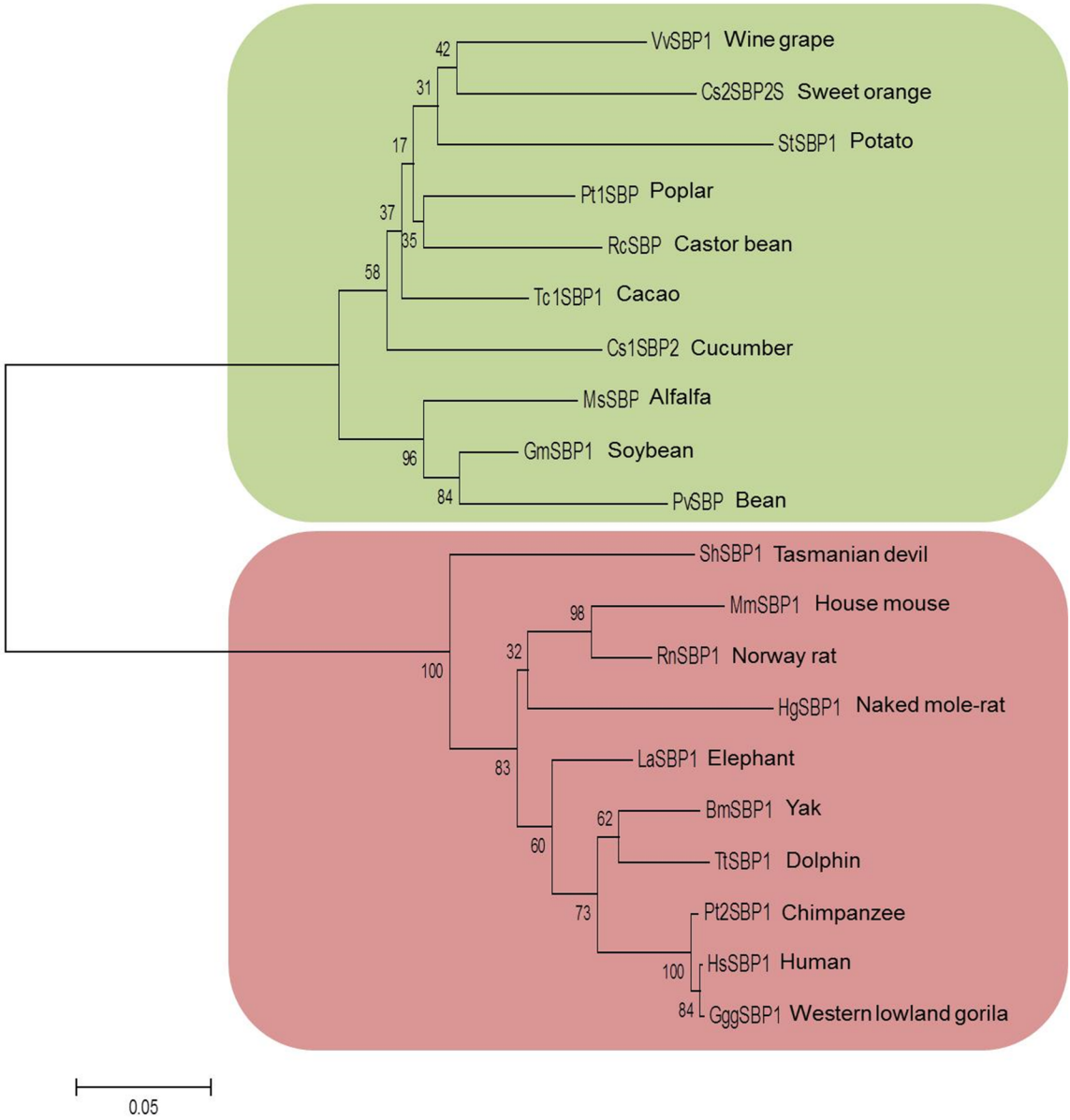
615 *Homo sapiens* (AAB02395.1; GI:1374792); GggSBP, *Gorilla gorilla*
616 *gorilla* (XP_004026682.1XP_004026682.1; GI: 426331426).

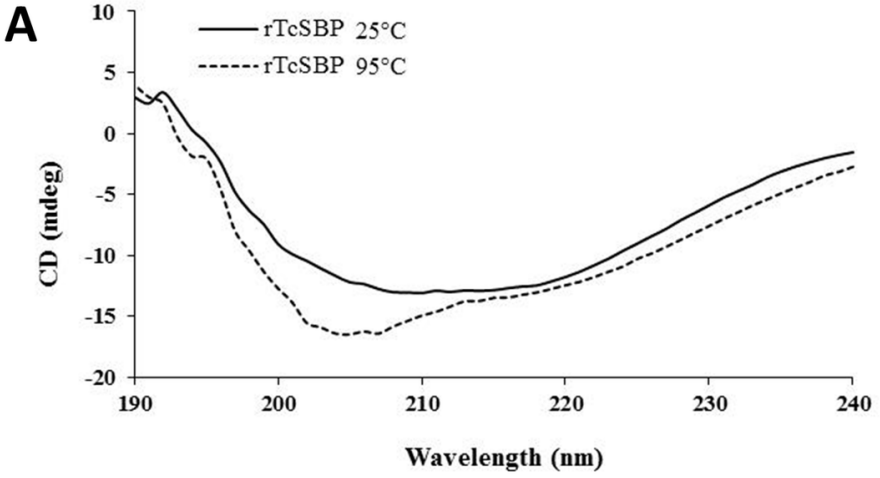
617 **Figure 4.** Circular dichroism analysis of rTcSBP. **A.** Spectra of the rTcSBP secondary
618 structure at 190-240 nm. Black and dashed lines corresponded to the rTcSBP structure
619 at 25°C and 95°C, respectively. **B.** Unfolding (black line) and refolding (dashed line) of
620 rTcSBP in phosphate buffer (pH 7.0) at 205 nm. **C.** Secondary structure of rTcSBP in
621 presence of sodium selenite at 25ºC. The grey line corresponds to sodium phosphate
622 buffer scan, yellow line to sodium selenite scan, orange line to rTcSBP scan, and blue
623 line to rTcSBP with sodium selenite scan (1:1 ratio).

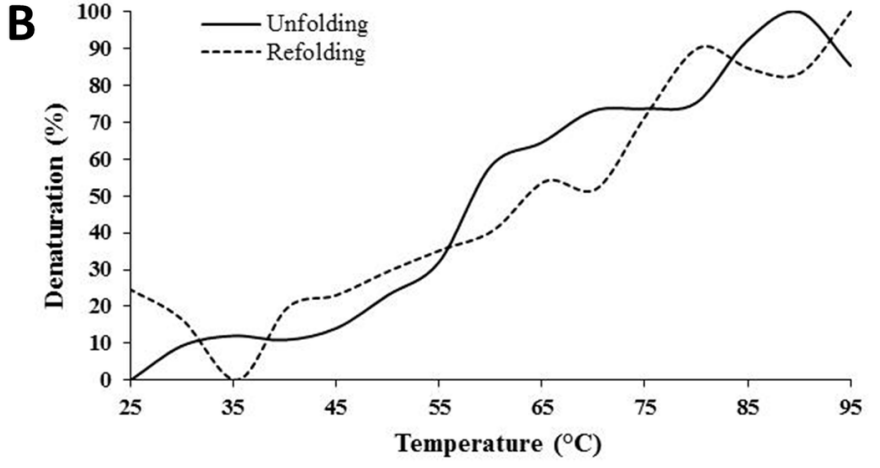
624 **Figure 45.** Relative expression of *TcSBP* in cacao meristems inoculated or not (control)
625 with *M. perniciosa*. **A.** Representation of the plant symptoms and fungus phase during
626 the infection time course in Catongo genotype. **B.** Relative expression observed in
627 resistant (TSH1188) and susceptible (Catongo) varieties. Different letters indicate
628 significant statistical difference between samples by the Scott-Knott test ($p \leq 0.01$). *t*-test
629 were applied with a critical value of 0.01: upper case letters correspond to statistics
630 between each of the genes on different harvesting times for each genotype while lower
631 case letters correspond to statistics between the three genes for each harvesting time.
632 hai: hours after inoculation; dai: days after inoculation. (*) indicates the times that were
633 harvested also in the non-inoculated (control) plants.

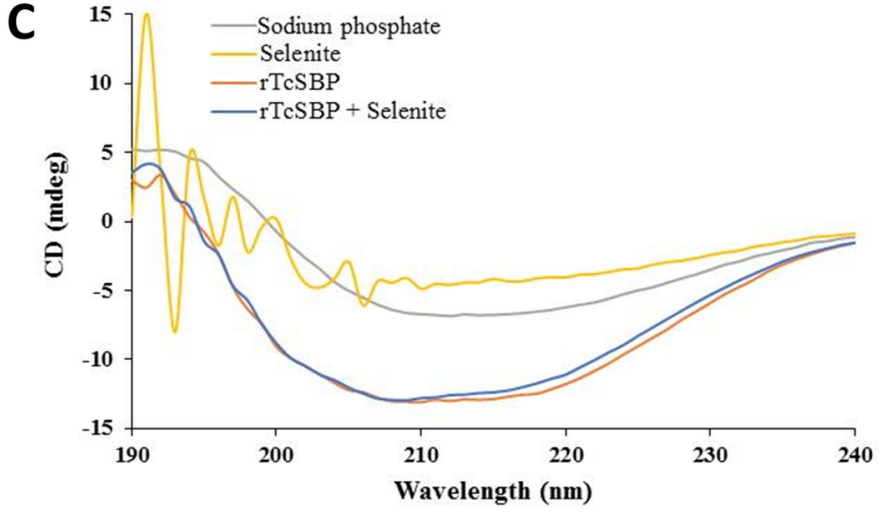

DNA
Chromosome 4
56
112
1063
196
859
228
153
232
150
108
146
273
514
282
402
4774 bp
Mature RNA
56
1431
402
1431 bp
Protein
TcSBP
476 aa
Untranslated region
Exons
Introns
Open Reading Frame

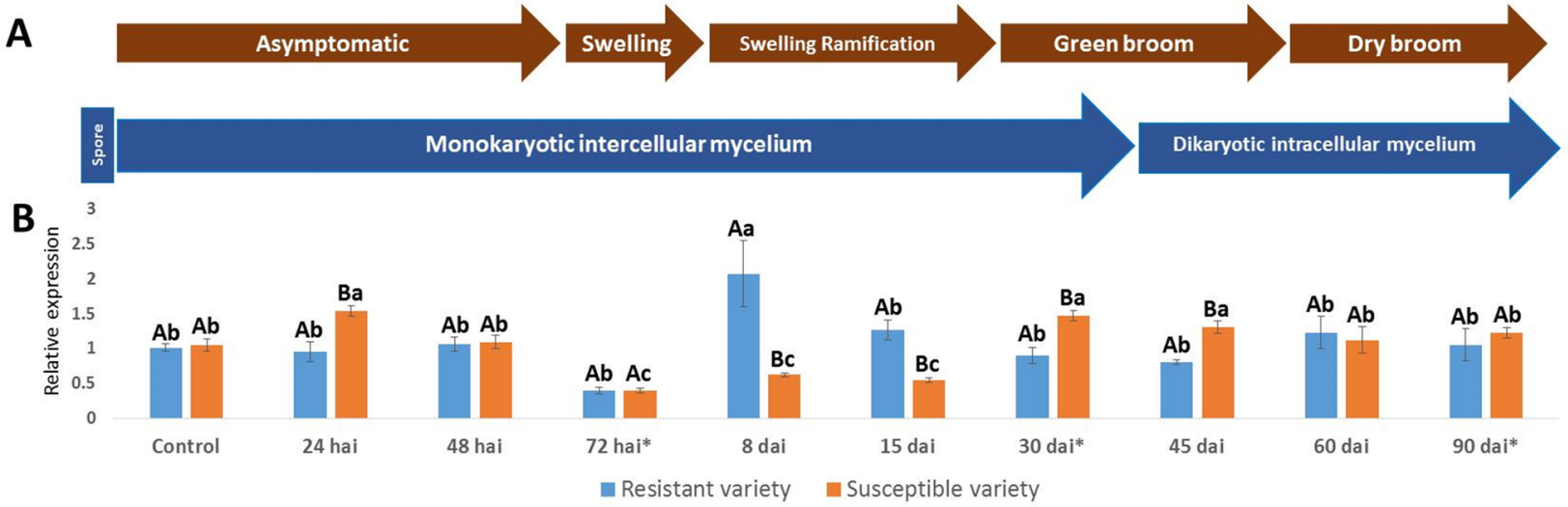

A
StSBP_Archaea
TcSBP_Plant
HsSBP1_Human
B
C
D
SER-85
SER-84
CYS-86
CYS-83
E


VvSBP1 Wine grape
Cs2SBP2S Sweet orange
StSBP1 Potato
Pt1SBP Poplar
RcSBP Castor bean
Tc1SBP1 Cacao
Cs1SBP2 Cucumber
MsSBP Alfalfa
GmSBP1 Soybean
PvSBP Bean
ShSBP1 Tasmanian devil
MmSBP1 House mouse
RnSBP1 Norway rat
HgSBP1 Naked mole-rat
LaSBP1 Elephant
BmSBP1 Yak
TtSBP1 Dolphin
Pt2SBP1 Chimpanzee
HsSBP1 Human
GggSBP1 Western lowland gorila
42
31
17
37
35
58
96
84
98
32
100
83
62
60
73
100
84
0.05


A
rTcSBP 25°C
rTcSBP 95°C
CD (mdeg)
10
5
0
-5
-10
-15
-20
190
200
210
220
230
240
Wavelength (nm)


B
Unfolding
Refolding
Denaturation (%)
100
90
80
70
60
50
40
30
20
10
0
25
35
45
55
65
75
85
95
Temperature (°C)


C
Sodium phosphate
Selenite
rTcSBP
rTcSBP + Selenite
CD (mdeg)
15
10
5
0
-5
-10
-15
190
200
210
220
230
240
Wavelength (nm)


A
Asymptomatic
Swelling
Swelling Ramification
Green broom
Dry broom
Spore
Monokaryotic intercellular mycelium
Dikaryotic intracellular mycelium
B
Relative expression
3
2.5
2
1.5
1
0.5
0
Ab Ab
Ab Ba
Ab Ab
Ab Ac
Aa Bc
Ab Bc
Ab Ba
Ab Ba
Ab Ab
Ab Ab
Control
24 hai
48 hai
72 hai*
8 dai
15 dai
30 dai*
45 dai
60 dai
90 dai*
Resistant variety
Susceptible variety