MOLECULAR PLANT PATHOLOGY (2007) **8**(3), 279–292 DOI: 10.1111/J.1364-3703.2007.00393.X

Genes differentially expressed in *Theobroma cacao* associated with resistance to witches' broom disease caused by *Crinipellis perniciosa*

GILDEMBERG AMORIM LEAL JR[1,2], PAULO S. B. ALBUQUERQUE[3] AND ANTONIO FIGUEIRA[1,*]

[1]*Centro de Energia Nuclear na Agricultura, Universidade de São Paulo; Ave. Centenário, 303, CP 96, Piracicaba, SP, 13400-970, Brazil*
[2]*Escola Superior de Agricultura 'Luiz de Queiroz', Universidade de São Paulo, Ave. Pádua Dias, 11, CP 83, Piracicaba, 13418-900, SP, Brazil*
[3]*ESJOH, Comissão Executiva do Plano da Lavoura Cacaueira, BR 316 km 17, CP 46, Marituba, 67105-970, PA, Brazil*

SUMMARY

The basidiomycete *Crinipellis perniciosa* is the causal agent of witches' broom disease of *Theobroma cacao* (cocoa). Hypertrophic growth of infected buds ('brooms') is the most dramatic symptom, but the main economic losses derive from pod infection. To identify cocoa genes differentially expressed during the early stages of infection, two cDNA libraries were constructed using the suppression subtractive hybridization (SSH) approach. Subtraction hybridization was conducted between cDNAs from infected shoot-tips of the susceptible genotype 'ICS 39' and the resistant 'CAB 214', in both directions. A total of 187 unique sequences were obtained, with 83 from the library enriched for the susceptible 'ICS 39' sequences, and 104 for the resistant 'CAB 214'. By homology search and ontology analyses, the identified sequences were mainly putatively categorized as belonging to 'signal transduction', 'response to biotic and abiotic stress', 'metabolism', 'RNA and DNA metabolism', 'protein metabolism' and 'cellular maintenance' classes. Quantitative reverse transcription amplification (RT-qPCR) of 23 transcripts identified as differentially expressed between genotypes revealed distinct kinetics of gene up-regulation at the asymptomatic stage of the disease. Expression induction in the susceptible 'ICS 39' in response to *C. perniciosa* was delayed and limited, while in 'CAB 214' there was a quicker and more intense reaction, with two peaks of gene induction at 48 and 120 h after inoculation, corresponding to morphological and biochemical changes previously described during colonization. Similar differences in gene induction were validated for another resistant genotype ('CAB 208') in an independent experiment. Validation of these genes corroborated similar hypothetical mechanisms of resistance described in other pathosystems.

* *Correspondence*: Tel.: +55 19 3429 4814; Fax: +55 19 3429 4610;
E-mail: figueira@cena.usp.br

INTRODUCTION

Witches' broom, caused by the basidiomycete *Crinipellis perniciosa* (Stahel) Singer (Tricholomataceae), is a severe disease of *Theobroma cacao* L. (cocoa), restricted to South America and the Caribbean (Purdy and Schimdt, 1996). The disease is endemic to the Amazon region, and upon introduction in producing areas, it has been responsible for the collapse of the cocoa industry in Surinam, Trinidad, Ecuador and more recently in Brazil (Andebrhan *et al.*, 1999). Estimated losses from witches' broom reach over 450 000 tonnes per year in American producing countries (Bowers *et al.*, 2001).

Basidiospores infect meristematic tissues (shoots, flower cushions, single flowers and developing pods), inducing a range of symptoms depending on organ infected and stage of development (Purdy and Schimdt, 1996). Hypertrophic growth of infected buds ('brooms') is the most dramatic symptom. Flower cushion infection usually leads to the production of vegetative shoots, abnormal flower development and parthenocarpic fruits. Pod infection can directly result in seed losses, whereas pod set is further reduced indirectly by the infection of flower cushions and the general debilitation of the tree (Andebrhan *et al.*, 1999). Under suitable conditions, infection is initiated by germ tubes of basidiospores penetrating stomata in meristematic tissue (Frias *et al.*, 1991). During penetration and intercellular colonization of tissues, no specialized structure is observed, and the pathogen is considered to be at the biotrophic phase of development, with monokaryotic hyphae (Frias *et al.*, 1991; Silva and Matsuoka, 1999). The first symptoms appear from 15 to 30 days after infection, with hyperplastic and hypertrophic growth of shoots. Later, tissues become necrotic and the pathogen becomes dikaryotic with clamp connections (Delgado and Cook, 1976). With the onset of the rainy season, basidiocarps emerge from necrotic tissues ('dry brooms'), producing basidiospores (Purdy and Schimdt, 1996).

Hyperplastic and hypertrophic growth of cortex cells have been described during the early stages of colonization (between 6 and 72 h after penetration), without the detection of fungal mycelia (Laker *et al.*, 1988; Silva and Matsuoka, 1999). Mycelia were more noticeable 7 days after infection (Frias *et al.*, 1991) and

tissue colonization occurred in both susceptible and resistant genotypes (Silva and Matsuoka, 1999). Colonization rates on differential genotypes have been related to isolate pathogenicity (Wheeler, 1985).

The ultimate method of witches' broom control will be through genetic resistance, as chemical and biological controls are inefficient, and phytosanitary pruning of infected tissues is a stop-gap measure (Purdy and Schimdt, 1996). Evaluation of resistance is based on incidence and/or severity of symptoms, and mechanisms of resistance are still largely unknown (Silva *et al.*, 1998). The genotype 'Scavina 6', originally collected in Peru, was the only recognized source of resistance, widely used in cocoa breeding programmes. 'Scavina 6' resistance appears to be dominant and predominantly monogenic (Faleiro *et al.*, 2006). However, the resistance from 'Scavina 6' has been overcome in Ecuador, and clones with this resistance gene have presented increasing infection symptoms in Brazil. Loss of resistance may be associated with the genetic variability of *C. perniciosa*, and new sources of resistance have been intensively searched for in the Brazilian Amazon (Sereno *et al.*, 2006). Among the novel sources of resistance, the accessions 'CAB 214' and 'CAB 208' have displayed significant levels of witches' broom resistance in southern Bahia and Amazonia, Brazil (Albuquerque and Figueira, 2004).

One approach to elucidating mechanisms of resistance and/or defence consists of the identification and characterization of differentially expressed genes between genotypes contrasting for response to infection by pathogens. In *T. cacao*, genes associated with resistance and defence response have been identified in expressed sequence tag (EST) collections derived from unchallenged leaves and seeds (Jones *et al.*, 2002), and from leaves treated with resistance elicitors, such as ethylene, methyl jasmonate and the fungal necrotic protein NEP1 (Verica *et al.*, 2004). The suppression subtractive hybridization (SSH) approach (Diatchenko *et al.*, 1996) is a more objective alternative to large-scale sequencing of EST libraries, because it is based on the subtraction of cDNAs from tissues and/or genotypes under distinct biological conditions to enrich the library for differentially expressed genes in quantitative and qualitative terms. SSH libraries have been used to identify induced genes in plant–pathogen interactions, such as *Pisolithus tinctorios* × *Eucalyptus* (Voiblet *et al.*, 2001) and *Puccinia triticina* × wheat (Thara *et al.*, 2003). Other studies applying SSH included the interactions between *Verticillium dahliae* and cotton (Zuo *et al.*, 2005), *Magnaporthe grisea* and rice (Lu *et al.*, 2004), *Gaeumannomyces graminis* and wheat (Guilleroux and Osbourn, 2004), and *Hemileia vastatrix* and coffee (Fernandez *et al.*, 2004). SSH libraries have also been used to identify induced genes in symbiotic interaction between the mycorrhizal fungus *Laccaria bicolor* and *Pinus* (Podila *et al.*, 2002).

The objective of this work was to identify induced genes of *T. cacao* in response to inoculation by *C. perniciosa* during the early asymptomatic stage of the witches' broom disease based on two SSH libraries, subtracted in both directions, contrasting cDNA derived from shoot tips of an inoculated susceptible genotype ('ICS 39') and a resistant one ('CAB 214'), sampling the transcriptome at 24, 48, 72, 120 and 240 h after inoculation.

RESULTS

Inoculation efficiency and evaluation of symptoms

A sample of five inoculated plants of each of the genotypes 'ICS 39', 'CAB 214' and 'CAB 208' plus 45 seedlings from the susceptible control 'Catongo' were evaluated for symptoms 30 days after inoculation to confirm if inoculation was successful. At least one type of witches' broom symptom had developed in 82% (37 out of 45 seedlings) of 'Catongo' seedlings by this time. Among the five inoculated 'ICS 39' plants, four were positive for the disease, while among the five plants for each resistant genotype ('CAB 214' or 'CAB 208'), only one of each presented the minor reaction of stem swelling, which is not considered a susceptible reaction.

Confirmation of the presence of *C. perniciosa* in inoculated cocoa plants

Detection of *C. perniciosa* in inoculated tissues was based on typical dissociation curves of fungal-specific *RpL35* amplicons from RT-PCR, using cDNA from *in vitro* cultivated mycelia as the positive control. The amplicon with identical dissociation temperature to the positive control was detected only in inoculated samples, but never in amplifications from non-inoculated control samples (Fig. 1A–C). Detection of *C. perniciosa RpL35* transcripts from inoculated 'ICS 39' samples was possible at 24, 48, 72, 120 and 240 h after inoculation based on melting curve analyses (Fig. 1A). Amplified fragments with the expected size were also detected by gel electrophoresis (Fig. 1D–F). For the resistant genotype 'CAB 214', *RpL35* transcripts were detected only at 24 and 48 h by melting curve analysis (Fig. 1B), but only at 24 h by gel analysis (Fig. 1E). For 'CAB 208' samples, the specific amplicon was detected later, only at 48, 72 and 240 h after inoculation when analysed by melting curve analysis (Fig. 1C) or by gel electrophoresis (Fig. 1F).

Construction of SSH libraries from inoculated *T. cacao*

Two SSH libraries enriched for sequences differentially expressed during the *T. cacao* and *C. perniciosa* interaction were generated for 'ICS 39' and 'CAB 214'. One SSH library was constructed using as tester pooled cDNA from all sampling periods after inoculation of the susceptible genotype ('ICS 39'), subtracted with a similar driver derived from the inoculated resistant genotype ('CAB 214'). The second SSH library was enriched for genes more active in the resistant 'CAB 214', because the driver and tester were inverted in comparison with the previous library.

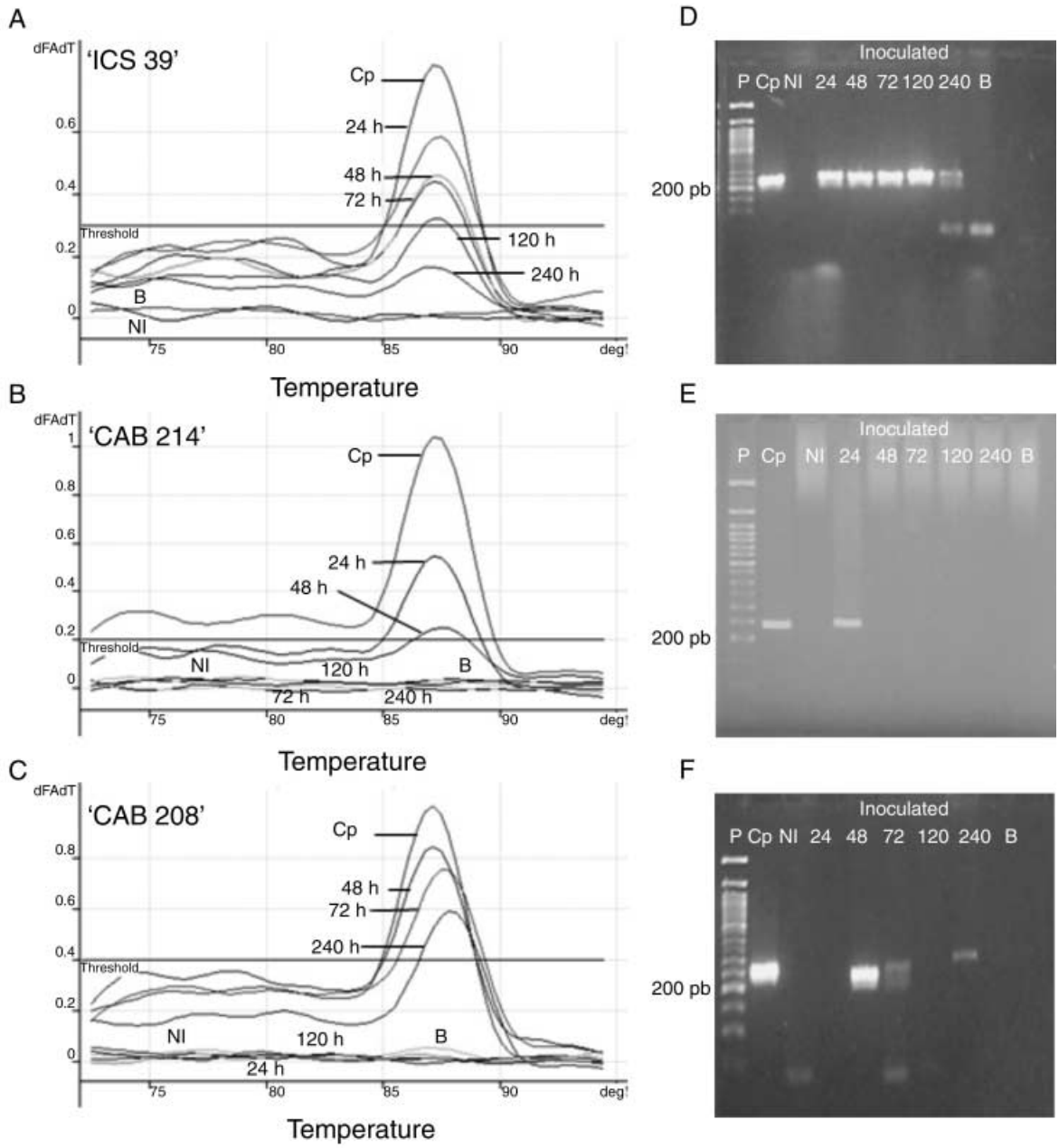

Fig. 1 Dissociation curve of the *RpL35* amplicon used to detect *C. perniciosa* in controls and inoculated plants from the genotypes 'ICS 39' (A), 'CAB 214' (B) and 'CAB 208' (C), and visualization of amplification products in 1.5% agarose gel electrophoresis for the same treatments, 'ICS 39' (D), 'CAB 214' (E) and 'CAB 208' (F). Treatments include positive controls using cDNA derived from *in vitro* cultivated mycelia (Cp); negative controls represent non-inoculated plant (NI) and a reaction without cDNA (B). Samples collected at 24, 48, 72, 120 and 240 h after inoculation for the genotypes 'ICS 39', 'CAB 214' and 'CAB 208' were evaluated by RT-qPCR. P = 50-bp ladder molecular weight standard.

A total of 960 clones were obtained, 430 from each library. Randomly, a total of 288 clones were sequenced from the enriched library for 'ICS 39' sequences, while 319 were sequenced from 'CAB 214', with an average size of *c.* 400 bp, ranging from 200 to 900 bp. Sequences from each library were individually clustered into contigs by CAP3 (Huang and Madan, 1999). The 288 clones from the 'ICS 39' enriched library were assembled into 32 contigs, derived from 114 sequences, while 174 remained as singletons, representing 206 unique sequences (unigenes), with an estimated level of redundancy of 28.5%. For the 'CAB 214' SSH library, the 319 sequenced clones were clustered into 29 contigs, from 103 sequences, leaving 216 singletons, in a total of 245 unigenes, with

an estimated redundancy of 23.2%. The low level of redundancy for both libraries indicated that further sequencing would probably reveal novel sequences.

Sequence annotation

Annotation of the 206 assembled sequences from the 'ICS 39' SSH library allowed the functional classification of 108 putative genes, with the remaining sequences classified either as 'no hit' or with *E*-values > 10^{-3} (see supplementary Table S1). From these 108 sequences, it was possible to identify putative functions to 83, including hypothetical and unknown proteins, leaving 25 that matched ESTs without annotation. These 108 functionally classified sequences were also compared with the *T. cacao* gene index database from The Institute for Genomic Research (TIGR), and 30 presented positive matches, including 27 with known functions, while the other three were not assigned to known functions. From the 30 sequences with positive identification to the *T. cacao* gene index, eight exclusively matched transcripts from the library of induced defence genes (Verica *et al.*, 2004), 15 matched members exclusive from the leaf and seed EST library (Jones *et al.*, 2002), while seven matched transcripts shared between both libraries.

In the 'CAB 214' SSH library, from the 245 assembled sequences, 127 presented matches with the databases analysed. From these 127 unique sequences, it was possible to identify putative functions to 104, including hypothetical and unknown proteins, leaving 23 that matched ESTs without annotation. These 127 sequences were also compared with the *T. cacao* gene index, and 51 presented positive matches, including 46 with known functions. From these 51 sequences with positive identification to the *T. cacao* gene index, 15 were exclusively matched to transcripts from the library of induced defence genes (Verica *et al.*, 2004), 26 matched members exclusive to the leaf and seed EST library (Jones *et al.*, 2002), whereas ten matched transcripts shared between both libraries.

The total of 187 annotated unigenes (83 from the 'ICS 39' and 104 from the 'CAB 214' library) were categorized by matching them with homologues from the Arabidopsis genome into the following functional classes (Fig. 2A): 'Response to stress and/or induced by abiotic/biotic stimuli', 'Energy metabolism', 'Metabolism', 'Unknown biological processes', 'Protein metabolism', 'Cellular organization and biogenesis', 'Signal transduction', 'DNA and RNA metabolism', and 'Cellular transport' (Fig. 2A). In this overall categorization, the class 'Metabolism' contained 19.3% of the sequences (36 out of 187), while the 'Response to stress and/or induced by abiotic/biotic stimuli' represented 18.2% (Fig. 2A). The classes 'Unknown biological processes' (12.9%), 'Energy metabolism' (11.2%), 'Protein metabolism' (14.4%), and 'DNA and RNA metabolism' (9.6%) presented a similar distribution, while the categories 'Cellular organization and biogenesis' (5.9%), 'Signal transduction' (6.4%), and 'Cellular transport' (2.1%) were less represented (Fig. 2A).

When the same categorization was performed for each SSH library individually, the 'ICS 39' enriched library displayed more sequences associated with categories of 'Response to stress and/or induced by abiotic/biotic stimuli' (25.3%, 21 out of 83; Fig. 2B); 'Energy metabolism' (13.3%) and 'Unknown biological processes' (16.9%). Conversely, the SSH library enriched for the resistant genotype presented more sequence representation of the categories 'Metabolism' (20.2%, 21 from 104; Fig. 2C); 'Protein metabolism' (18.3%), 'DNA and RNA metabolism' (13.5%), 'Cellular organization and biogenesis' (7.7%) and 'Cellular transport' (1.9%) (Fig. 2C).

Validation of libraries by quantitative amplification of reversed transcripts (RT-qPCR)

Twenty-three genes identified in the two *T. cacao* SSH libraries were chosen to validate the libraries, with 13 derived from the 'ICS 39' (susceptible) enriched library and ten from the 'CAB 214' (resistant) (Table 1). The transcripts were chosen to represent putative gene categories associated with resistance and/or defence mechanism (see supplementary Tables S1 and S2), including three sequences coding for 'leucine-rich proteins'—LRP (LRP-ICS, LRP-CAB, LRP.2-CAB) associated with membrane receptors; one putative gene coding for 'WD-40' (WD.40-ICS), related to signal transduction and control of growth by plant hormones; transcription factors, including 'ethylene responsive factor-type'—ERF (ERF-ICS), 'no apical meristem'—NAM (NAM-ICS) and 'short vegetative phase'—SVP (SVP-ICS); proteins related to pathogenesis and response to stress, including 'hevein' (Hev-ICS), 'thaumatin' (Thau-ICS), 'chitinase' (Chit-ICS), 'glucanase' (Gluc-CAB), 'cysteine proteinase' (Cyst-ICS) and 'peroxidases' (Perox-ICS; Perox-CAB); secondary metabolism enzymes, including 'caffeine synthase' (Caff-CAB), 'anthocyanin reductase' (AR-CAB) and 'leuco-anthocyanidin dioxygenase' (LAD-ICS), encoding key enzymes of tannin and anthocyanin biosynthesis; and 'epoxy-hydrolase' (EH-CAB), an enzyme associated with cutin biosynthesis. The other selected genes have been previously related to be induced in response to microrganism–plant interaction, including EIG.7 (EIG.7-CAB; Takemoto *et al.*, 2001), cytochrome oxidase P450 (C.P450-ICS), cellulase (Cell-CAB), USP protein (USP-CAB; Hohnjec *et al.*, 2000) and ATS3 (ATS3-ICS; Lee *et al.*, 2006) (see Supplementary Table S2).

The differential accumulation of these transcripts between the susceptible and resistant genotypes at 24, 48, 72, 120, and 240 h after inoculation was investigated by RT-qPCR, using cocoa actin transcript as internal gene reference. Amplification efficiencies for all the primer-pairs tested were around 2, with correlation coefficients above 0.99 for the dilution curve analyses (data not shown). Because amplification efficiencies were close to 2 and similar among transcripts, the $2^{-(\Delta\Delta Ct)}$ method (Livak and Schmittgen, 2001) was used to express relative gene expression differences between conditions.

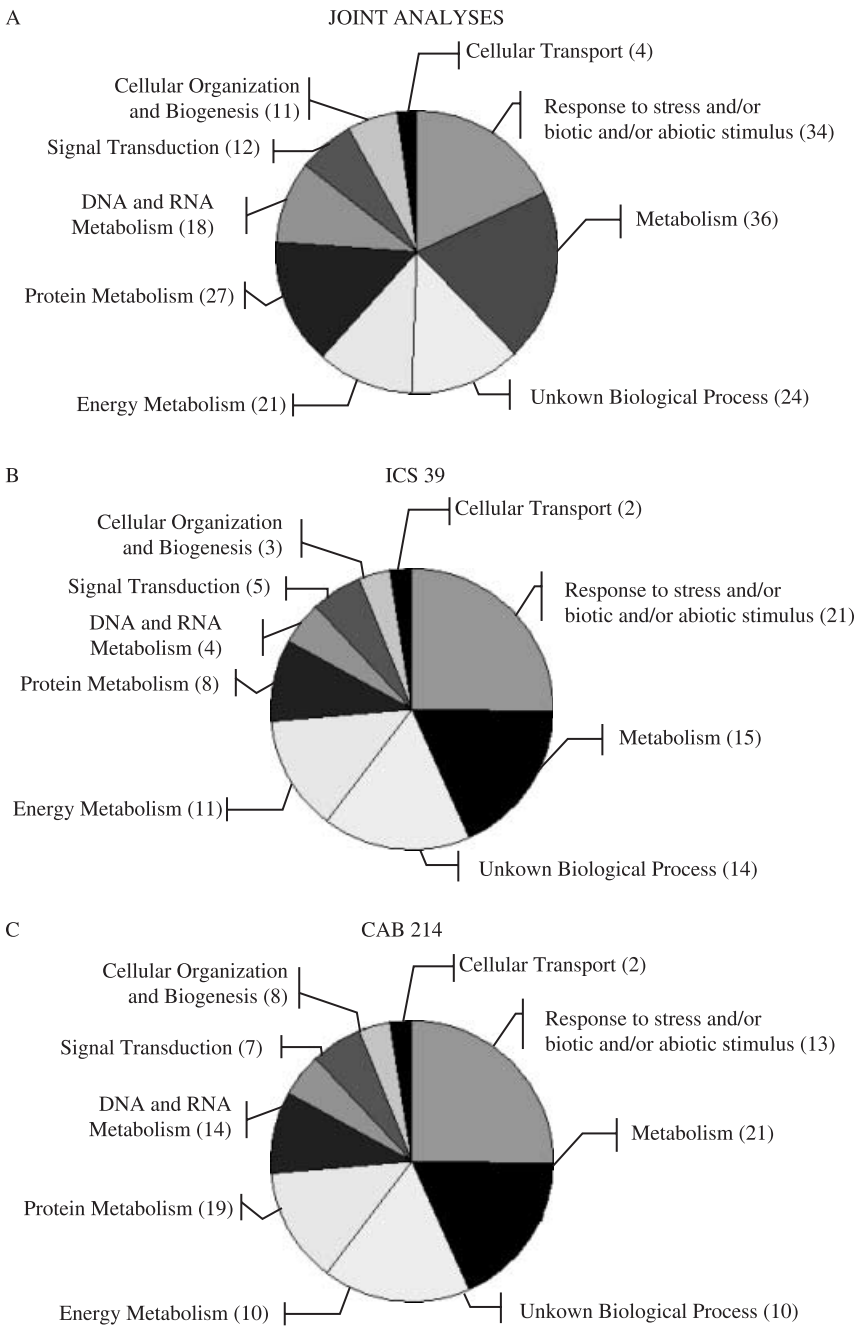

Fig. 2 Distribution for all unique sequences annotated for susceptible and resistant *T. cacao* genotypes inoculated with *C. perniciosa* into putative functional categories with respective number of members in parentheses (A). Putative functional categories for unique sequences identified for each genotype: susceptible 'ICS 39' (B) and resistant 'CAB 214' (C).

Initially, the level of gene expression was estimated for non-inoculated control samples from both genotypes for all periods (24–240 h) to establish an average baseline expression reference, without pathogen interaction. In general, the relative presence of all 23 transcripts in the non-inoculated susceptible and resistant genotypes was nearly constant over all five sampling periods (data not shown), except for 24 h, which appeared to present a peak of transcription for some of the genes (Fig. 3A), possibly due to mechanical induction during the mock inoculation process. Thus, a relative mean value of gene expression for each transcript over the five sampling periods was obtained for each genotype.

To determine differences in basal gene expression at 24 h in relation to the overall average of transcript presence, an arbitrary threshold value of 2.8 (Benitez *et al.*, 2005) was adopted to consider significant differences (Fig. 3A). In the non-inoculated susceptible 'ICS 39', only Caff-ICS and Hev-ICS presented significant expression above average at 24 h (Fig. 3A). For the non-inoculated 'CAB 214', there was a general trend for a differential

Table 1 Genes selected to be validated for expression by quantitative reverse transcription amplification, with gene abbreviation, full name of gene and function of gene.

Gene abbreviation*	Full name of gene	Function of gene
LRP-ICS	Leucine-rich protein	Membrane receptor
LRP-CAB	Leucine-rich protein	Membrane receptor
LRP.2-CAB	Leucine-rich protein	Membrane receptor
WD.40-ICS	WD-repeat protein	Signal transduction
ERF-ICS	Ethylene-Response Factor 1 protein	Transcription regulator
NAM-ICS	No-Apical Meristem protein	Transcription regulator
SVP-ICS	Short Vegetative Phase protein	Transcription regulator
Hev-ICS	Hevein	A pathogenesis related protein
Thau-ICS	Thaumatin	A pathogenesis related protein
Chit-ICS	Chitinase	Hydrolysis of chitin
Gluc-CAB	Beta 1,3 glucanase	Hydrolysis of glucan
Cyst-ICS	Cysteine Proteinase	Peptidase
Perox-ICS	Peroxidase	Catalyse oxidative reaction
Perox-CAB	Peroxidase	Catalyse oxidative reaction
Caff-CAB	Caffeine synthase	Biosynthesis of caffeine
AR-CAB	Anthocyanidin reductase	Biosynthesis of tannin and pro-anthocianydins
LD-ICS	Leucoanthocyanidin dioxygenase	Biosynthesis of anthocyanidin
EH-CAB	Epoxide hydrolase	Biosynthesis of cutin
C.P450-ICS	Cytochrome P450	Oxidative degradation
Cell-CAB	Endo-1 4-beta-glucanase	Hydrolysis of cellulose
USP-CAB	Universal Stress Proteint	Response to stress
EIG.7-CAB	Lipid-associated family protein	Unknown
ATS.3-ICS	Embryo-specific protein	Unknown

*The suffixes ICS and CAB refer to genes from the library of the susceptible genotype ICS 39 and the resistant genotype CAB 214, respectively.

higher frequency only at 24 h, with higher levels for Thau-ICS, Hev-ICS, NAM-ICS, LRP-ICS, C.P450-ICS, LRP.2-CAB and Perox-ICS (Fig. 3A). Next, important differences in basal level of expression between genotypes were estimated for all genes using the susceptible 'ICS 39' as reference (Fig. 3B), considering the significant threshold limit of 2.8 (Benitez *et al.*, 2005). Significant differences in transcript accumulation in non-inoculated 'CAB 214' in comparison with 'ICS 39' were detected only at 24 h for six of the same genes (Thau-ICS, NAM-ICS, C.P450-ICS, LRP-ICS, LRP.2-CAB and Perox-ICS), except for Hev-CAB (Fig. 3B), which was more expressed in 'ICS 39'. Hev-CAB was also more highly expressed in 'ICS 39' at 48, 72, 120 and 240 h after treatment (data not shown).

Subsequently, the increase in transcript accumulation in the inoculated samples for each genotype ('ICS 39', 'CAB 214', 'CAB 208') was evaluated using the non-inoculated samples for the same period (after 48 h) as reference (Fig. 4). For 'ICS 39', from the 23 genes analysed, 16 appeared to be induced in comparison with non-inoculated controls, generally at two points in time: an early peak at 48 and/or 72 h, and a late peak at 120 and/or 240 h after inoculation (Fig. 4). Ten transcripts presented both peaks of accumulation (LRP-ICS, LRP-CAB, LRP.2-CAB, WD.40-ICS, ERF-ICS, Hev-ICS, Thau-ICS, Cyst-ICS, EH-CAB and Cell-CAB), while four genes presented only the late induction at 120 and/or 240 h (NAM-ICS, Caff-CAB, EIG7-CAB and C.P450-ICS). Only two transcripts (AR-CAB and LAD-ICS) presented an early accumulation in 'ICS 39' at 48 and/or 72 h (Fig. 4). Seven genes did not appear to be induced in the susceptible genotype upon inoculation (SVP-ICS, Chit-ICS, Perox-ICS, Perox-CAB, Gluc-CAB, ATS3-ICS and USP-CAB).

Conversely, for the resistant genotype 'CAB 214', from the 23 transcripts evaluated by RT-qPCR, 21 were induced, with the exception of Chit-ICS and Gluc-CAB (Fig. 4). From the 21 significantly induced transcripts, 14 presented an early accumulation (48–72 h after inoculation) together with a late one (120–240 h). Seven transcripts (SVP-ICS, Perox-ICS, Perox-CAB, Caff-CAB, AR-CAB, LAD-ICS, C.P450-ICS and EIG7-CAB) appeared to present only the early accumulation. In general, the most important gene induction occurred after 48 h. When inoculated samples from the other resistant genotype ('CAB 208') were evaluated, from the same 23 transcripts analysed, 21 appeared to accumulate in inoculated samples, with only two (AR-CAB and C.P450-ICS) without displaying any increase in transcripts (Fig. 4). A pattern of expression similar to 'CAB 214' was observed for 'CAB 208', with an early (48–72 h after inoculation) and a late peak of transcript accumulation (120–240 h). From the 21 transcripts with increased accumulation, 11 (LRP-ICS, LRP-CAB, NAM-ICS, SVP-ICS, Cyst-ICS, Gluc-CAB, EH-CAB, EIG7-CAB, Cell-CAB, ATS3-ICS and USP-CAB) exhibited both periods of induction, while eight presented only an early accumulation (LRP.2-CAB, WD.40-ICS,

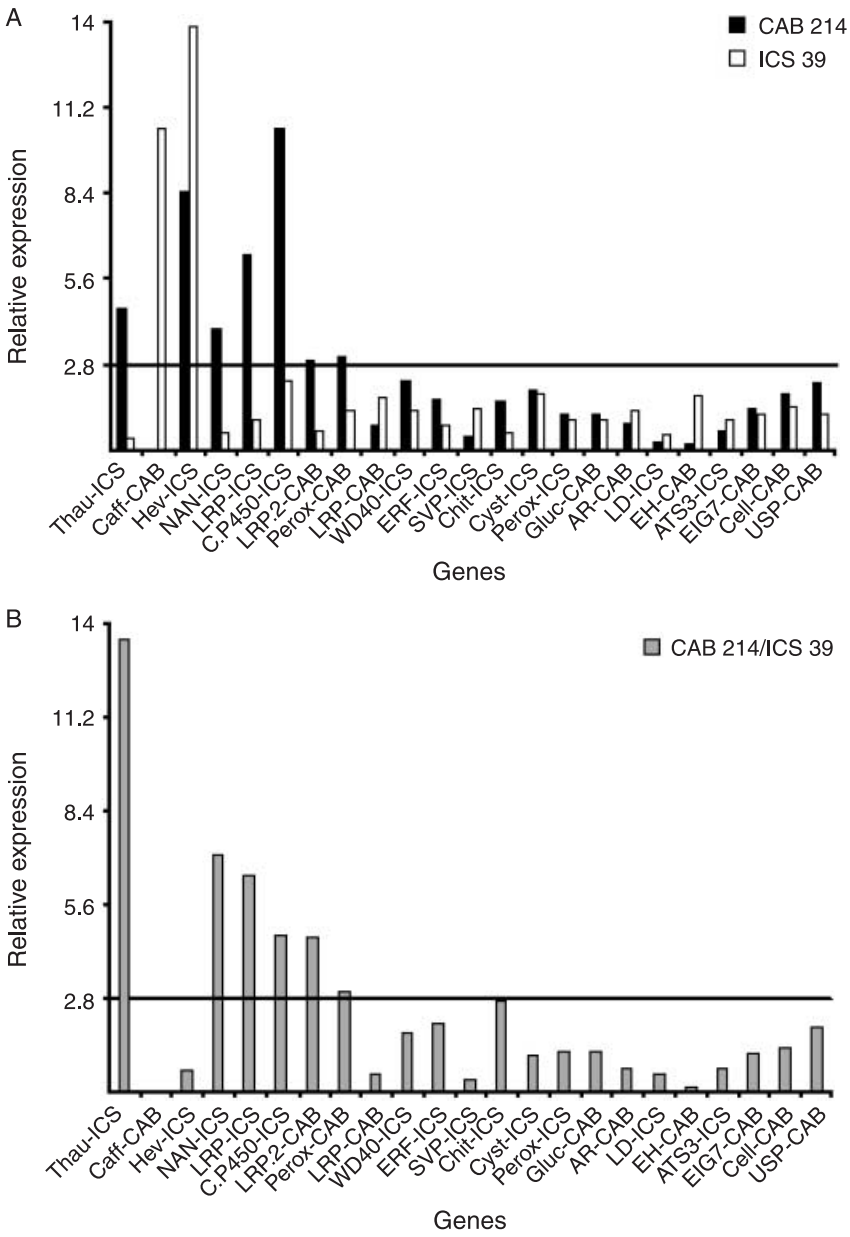

Fig. 3 Relative level of transcripts for 23 genes in non-inoculated *T. cacao* shoot tips from genotypes 'ICS 39' and 'CAB 214' at 24 h after treatment (A) in comparison with overall average of expression in five sampling periods (24, 48, 72, 120 and 240 h). Relative differential expression in 'CAB 214' in relation to 'ICS 39' at 24 h (B). The values were estimated based on mean C_T from three replicates.

ERF-ICS, Hev-ICS, Thau-ICS, Chit-ICS, Perox-ICS and Perox-CAB) and two the late accumulation (Caff-CAB and LAD-ICS). The transcripts and expression pattern from 'CAB 208' were distinct from 'CAB 214', with a remarkable higher induction of ten genes at 72 h after inoculation (Fig. 4).

DISCUSSION

In this study, the transcriptomes of infected shoot tips of the cocoa witches' broom-resistant ('CAB 214') and susceptible genotypes ('ICS 39') were compared by SSH analyses. We successfully identified genes whose expression might be associated with early resistance mechanisms of *T. cacao* plants against *C. perniciosa* infection. We chose to perform the subtraction hybridization confronting cDNA from infected tissues from genotypes contrasting for resistance aiming to identify genes specifically induced in each genotype under pathogen challenge during the asymptomatic stage. During this early stage of infection, morphological alterations, such as hyperplastic and hypertrophic growth of cortex cells (Silva and Matsuoka, 1999), as well as biochemical changes, including the variation in the fluorescence of host cells (Frias *et al.*, 1991) and the synthesis of tannins (Scarpari *et al.*, 2005), have been described.

The detection of the ribosomal protein gene *RpL35* transcripts from *C. perniciosa* enabled the detection of the pathogen in susceptible plants for all sampling periods (Fig. 1), corroborating the data from symptom evaluation. Molecular detection of

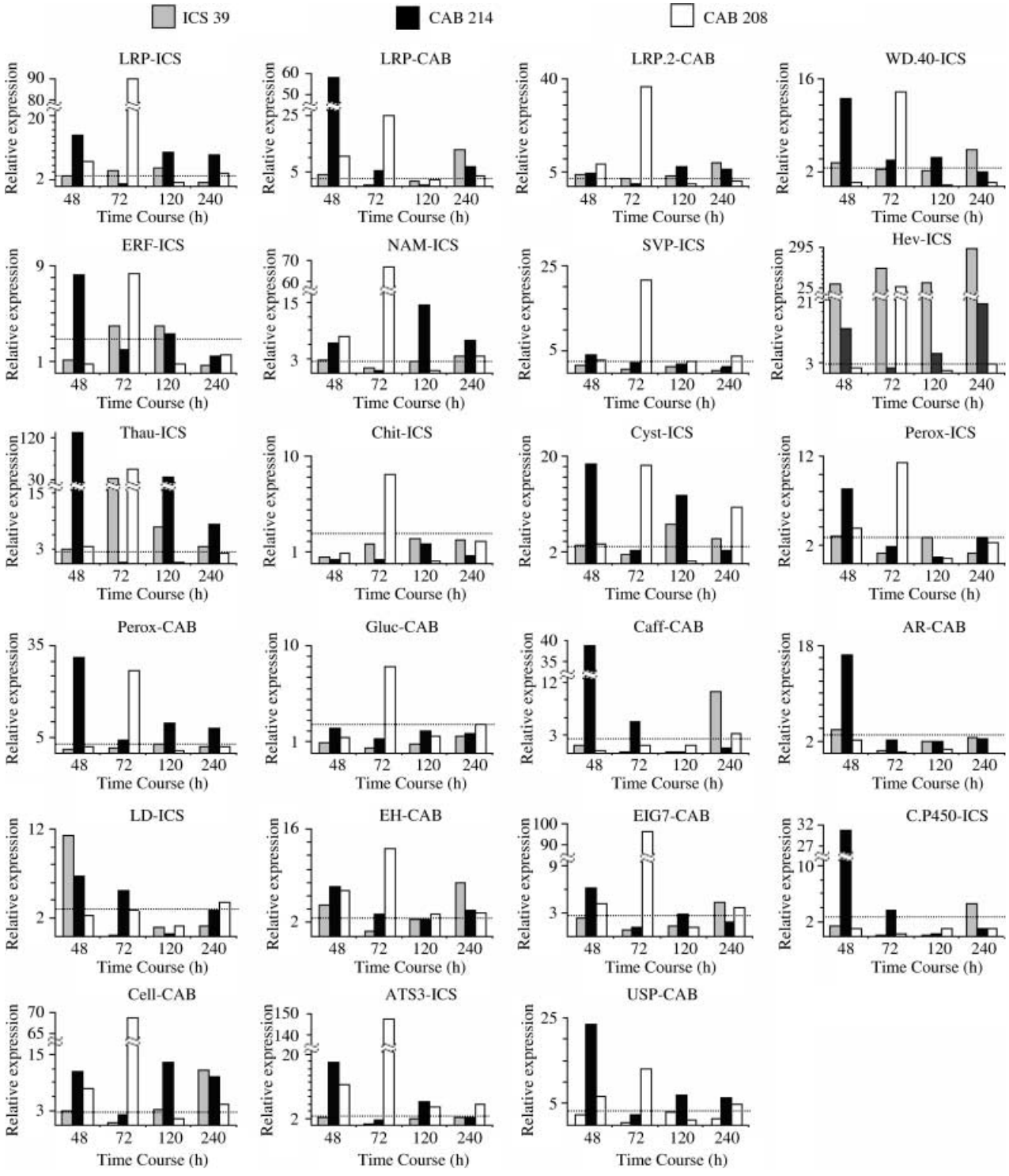

Fig. 4 Relative induction of 23 genes putatively associated with resistance and/or defence mechanisms in shoot-tip samples from three genotypes ('ICS 39', 'CAB 214' and 'CAB 208') inoculated with *Crinipellis perniciosa* in relation to non-inoculated controls. Expression values were estimated based on quantitative RT-PCR using the $2^{-\Delta(\Delta Ct)}$ method (Livak and Schmittgen, 2001), at four sampling periods (48, 72, 120 and 240 h after inoculation) for genes identified in suppression subtractive libraries enriched for sequences from the susceptible genotypes (–ICS) or from the resistant (–CAB). The arbitrary significance threshold value of 2.8 (Benitez *et al.*, 2005) is shown as a line across genotypes and time.

C. perniciosa in the resistant genotypes 'CAB 214' and 'CAB 208' confirmed the successful inoculation (Fig. 1), but also corroborated the apparent restriction of infection by these two genotypes, limiting the presence of detectable mycelia. The pattern of *C. perniciosa* detection on 'CAB 214' differed from 'CAB 208' possibly due to an earlier (less than 48 h) restriction of colonization in 'CAB 214', in contrast to the later detection in 'CAB 208' (Fig. 1). Limited colonization of infected tissues in resistant genotypes has been reported based on mycelia detection by optical or electron microscopy (Silva and Matsuoka, 1999; Wheeler, 1985). Therefore, the resistant genotypes analysed here appeared to differ for rate of fungal colonization, despite the fact that both share similar resistant phenotypes. 'CAB 208' seemed to be less efficient in early restriction of infection, which might explain the higher induction of resistance and/or defence mechanism genes at 72 h (Fig. 4). Further experiments would be required to confirm the use of *RpL35* as an early detection method for the presence of the pathogen, and for evaluation of cocoa genotype response.

The categorization of the 'ICS 39' enriched library (Fig. 2B) allowed the detection of an apparently larger proportion of transcripts associated with plant defence responses in comparison with that from the resistant library (Fig. 2C). By contrast, transcripts associated with signalling processes (including putative LRR proteins) were more frequent in the 'CAB 214' library ('Signal transduction'; Fig. 2C). However, no constitutive expression differences were detected for a sample of these transcripts (LRP-CAB, LRP.2-CAB and LRP-ICS) in non-inoculated samples, which could justify the differences in resistance to *C. perniciosa*, in contrast to what has been described for the interactions *Venturia inaequalis* and apple (Degenhardt *et al*., 2004), and *Hemileia vastatrix* and coffee (Fernandez *et al*., 2004). Similarly, no constitutive difference in expression of stress- and defence-related proteins was detected (Fig. 3).

The 'ICS 39' enriched library (Fig. 2B) also displayed a smaller number of transcripts classified as 'Metabolism', 'Protein metabolism', 'DNA and RNA metabolism', and 'Cellular organization and biogenesis', suggesting a down-regulation of the general metabolism under a host–pathogen compatible interaction. Similar responses have been described in other pathosystems, with repression of genes related to general metabolism (Logemann *et al*., 1995; Somssich and Hahlbrock, 1998). Conversely, in the resistant genotype 'CAB 214', there was a large representation of transcripts categorized into metabolism-related classes (Fig. 2C), especially those encoding enzymes associated with secondary metabolism related to phytoalexin synthesis.

The level of transcripts of 23 identified putative genes was found to be generally stable between sampling periods in non-inoculated plants from the three genotypes, indicating that most of these genes were only induced by pathogen challenge. A few transcripts that appeared to accumulate in non-inoculated resistant 'CAB 214' plants were probably affected by mechanical, wounding and/or abiotic stimuli from the mock-inoculation. No important differences in basal gene expression were detected between non-inoculated 'ICS 39' and 'CAB 214'. Differential constitutive expression between genotypes could indicate a gene signature, which could be associated with resistance or susceptibility (Kazan *et al*., 2001), and useful as a marker for assisted breeding (Holter *et al*., 2000). One exception was the increased accumulation in transcripts of the putative Hevein gene in 'ICS 39' (Hev-ICS; Fig. 3), homologous to the *Triticum monococcum* Hevein gene, encoding a peptide which binds to chitin from fungal cell walls (Bertini *et al*., 2006). Unfortunately, the increase in transcription of the putative Hevein gene was associated with susceptibility rather than resistance.

Most of the 23 genes seemed to be up-regulated by pathogen inoculation in the evaluated genotypes, but differed for induction kinetics. The resistant 'CAB 214' and 'CAB 208' displayed a stronger induction at 48 and 72 h for some of the genes evaluated, while the susceptible 'ICS 39' mainly exhibited a peak of accumulation only at 120 and/or 240 h (Fig. 4) after inoculation. In other pathosystems, susceptible or resistant genotypes have commonly differed quantitatively for gene expression, with the susceptible genotype displaying a later induction (Benitez *et al*., 2005; Tao *et al*., 2003), but not qualitatively. Here, qualitative differences for specific transcripts between genotypes were also observed, since from the 23 genes evaluated, only 16 were induced in the susceptible genotype, while 21 were induced in both resistant genotypes (Fig. 4). A second peak of up-regulation was also observed for the resistant genotypes, with equal or superior transcript accumulation in relation to the late induction of the susceptible genotype (Fig. 4). Two periods of gene up-regulation in inoculated plants were detected in the *Ustilago maydis—Zea mays* interaction, a first one during the initial asymptomatic stage, followed by another during symptom development (Basse, 2005). The cocoa samples collected at 240 h were closer to the onset of the first witches' broom symptoms (Delgado and Cook, 1976), suggesting a possible change in fungal metabolism, with the initiation of the necrotrophic phase, possibly eliciting a novel host defence response.

The search for resistance gene candidates identified those up-regulated in the resistant genotype, but not in the susceptible, including one transcription factor SVP (short vegetative phase), two peroxidases (Perox-ICS and Perox-CAB) and two genes with unknown functions (ATS-ICS and USP-CAB). The SVP-ICS shared similarity with a *Populus tomentosa* MAD-Box transcription factor, a plant-exclusive class regulating flower development and controlling flowering and vegetative growth (Ng and Yanofsky, 2001). The identified peroxidase homologues belonged to class III, with Perox-ICS as a putative homologue to a *Quercus suber* peroxidase (AAR31106.1) and Perox-CAB a homologue to a distinct peroxidase from *Vigna angularis* (BAA01950). The induction of peroxidase activity has been recognized as one of the first

events of a plant–pathogen interaction (Chittoor *et al.*, 1997), with roles in reinforcing cell walls and protecting against reactive oxygen species (Kawaoka *et al.*, 2003; Quiroga *et al.*, 2000). Peroxidases were identified in rice as defence responses against *Xanthomonas oryzae* pv. *oryzae* and *Magnaporthe grisea* (Chittoor *et al.*, 1997; Sasaki *et al.*, 2004). The ATS3-ICS sequences presented a PLAT domain ('polycystin-1-lipoxygenase-alpha-toxin') or LH2 ('lipoxygenase homology 2') domain, present in *Arabidopsis thaliana* seed-specific ATS3 (At2g41470) and in the *CAPIP2* gene of *Capsicum annuum*, involved in infection responses to *Xanthomonas campestris* and abiotic stress (Lee *et al.*, 2006). The USP-CAB gene shared homology with an unknown protein with a universal stress protein (USP) domain, reported to be induced in plants by ethylene (Sauter *et al.*, 2002) or at the final stage of nodulation in *Vicia faba* (Hohnjec *et al.*, 2000). USP proteins were first identified in the cytoplasm of bacteria, and later as stress response in plants (Kerk *et al.*, 2003).

Distinct up-regulation of AR-CAB and C.P450-ICS could discriminate the resistant genotype 'CAB 214' from 'CAB 208', while differences in transcript accumulation of the defence genes Chit-ICS and Gluc-CAB could distinguish 'CAB 208' from 'CAB 214' and 'ICS 39' (Fig. 4). The AR-CAB transcript appeared to be a putative homologue to a *Gossypium arboreum* gene coding for leuco-anthocyanin reductase, an enzyme that catalyses the biosynthesis of condensed tannins and pro-anthocianydins in cotton (Winkel-Shirley, 2001). Condensed tannins of *T. cacao* inhibited spore germination *in vitro* and biotrophic mycelia of *C. perniciosa* (Brownlee *et al.*, 1992). Furthermore, an increase in tannin biosynthesis was detected in susceptible plants of cocoa, 3 days after inoculation with *C. perniciosa* (Scarpari *et al.*, 2005). The induction of AR-CAB at the early stage of infection in 'CAB 214' supported a hypothesis on the importance of tannins during the potential initial containment of tissue colonization by the pathogen. The ICS-C.P450 sequence shared homology with a cytochrome monooxygenase P450 from *A. thaliana*. This gene family codes for enzymes involved in biosynthesis of secondary metabolites, growth regulators (gibberellin, jasmonate, brassinosteroids) and in cellular detoxification (Chapple, 1998). The Chit-ICS presented homology with a chitinase gene from *Gossypium hirsutum*, classified in Family 19 of the Glucanases, which contains chitinases from Class I, III and IV (Collinge *et al.*, 1993). The Gluc-ICS shared similarities with a glycosyl hydrolase gene from *A. thaliana*, similar to a β-1,3-glucanase from *Camellia sinensis* (GI:15150341). Chitinases and glucanases are recognized as enzymes associated with fungal cell wall degradation, restricting pathogenesis (Collinge *et al.*, 1993; Kauffmann *et al.*, 1987). Based on the differential expression for the identified defence-response gene candidates between both resistant genotypes, it could be hypothesized that 'CAB 214' appeared to inhibit *C. perniciosa* colonization by directing transcription for genes associated with synthesis of antimicrobial compounds, while 'CAB 208' tended to favour lytic enzymes that degrade fungal cell walls. Furthermore, the remarkable induction of resistance and/or defence mechanism genes at 72 h in 'CAB 208' might be related to a stronger response to a more intense colonization in comparison with 'CAB 214', which appeared to control infection at an earlier stage (before 48 h). Susceptibility appeared to arise from an ineffective late response. It is noteworthy that the qualitative and quantitative differential induction and kinetics between both resistant genotypes corroborated differences in the rate of colonization (Fig. 1) and the level of resistance, as 'CAB 214' displayed a lower incidence of witches' broom symptoms than 'CAB 208' in genetic studies (Albuquerque and Figueira, 2004). Furthermore, genetic mapping with preliminary quantitative trait loci (QTL) analyses suggested the existence of distinct resistance loci between the genotypes (Albuquerque and Figueira, unpublished data).

No transcript associated with signal transduction and transcription factors was identified that differed in expression between both resistant genotypes. To identify transcription factors associated with these witches' broom resistance sources, the promoter region of the up-regulated genes could be characterized. A good gene candidate for promoter analysis would be Caff-CAB, early induced in 'CAB 214', but later up-regulated in the other two genotypes (Fig. 4). Caffeine accumulated during the *T. cacao—C. perniciosa* interaction (Aneja and Gianfagna, 2001; Scarpari *et al.*, 2005), and caffeine presents *in vitro* activity against the pathogen (Aneja and Gianfagna, 2001). In *Coffea arabica*, a negative relationship between caffeine content and resistance was described (Guerreiro Filho and Mazzafera, 2000, 2003), but in *T. cacao*, this association could not be demonstrated because only plants with symptoms have been analysed (Aneja and Gianfagna, 2001; Scarpari *et al.*, 2005). The increase in caffeine might result from intensive transcriptional activity in response to biotic or abiotic stresses (Guerreiro Filho and Mazzafera, 2000). Induction of Caffeine synthetase by such distinct elicitors as benzothiadiazole or methyl jasmonate suggested that different forms of this enzyme might occur in *T. cacao* (Aneja and Gianfagna, 2001; Bailey *et al.*, 2005). An alternative function of this putative caffeine synthetase might involve a supposed methyl transferase activity, carrying out methylation of various compounds, including synthesis of methyl salicylate or methyl jasmonate (Ross *et al.*, 1999; Seo *et al.*, 2001), which might participate in plant defence responses.

The SSH approach was an efficient method to identify genes induced in response to *C. perniciosa* inoculation of *T. cacao*. Evaluation of transcript accumulation in response to pathogen inoculation revealed distinct kinetics of gene induction at the asymptomatic stage of the disease. Expression induction in the susceptible 'ICS 39' in response to *C. perniciosa* was delayed and limited, while in 'CAB 214', there was a quicker and more intense reaction, with two peaks of gene induction at 48 and 120 h after inoculation, corresponding to events already described during colonization. Similar differences in gene induction were validated

for another resistant genotype ('CAB 208') in an independent experiment. Some of the identified genes may be responsible for defence responses or may be resistance genes and their differential expression discriminated between the two resistant genotypes. The detection of gene polymorphism could be used to generate genetic markers associated with QTLs. The characterization of promoter regions could allow the identification of transcription factors involved with the differential level and timing of expression between genotypes.

EXPERIMENTAL PROCEDURES

Plant material and *C. perniciosa* inoculum

Clonal plants of the witches' broom-resistant genotypes 'CAB 214' and 'CAB 208, and the susceptible 'ICS 39' were obtained by budding on to common seedling rootstocks, while seedlings of the susceptible control 'Catongo' (an albino mutant) were derived from open-pollinated pods. All plants were grown for 2–3 months in the nursery at the 'Estação de Recursos Genéticos do Cacaueiro José Haroldo', Marituba, PA, Brazil (1°12′S, 49°13′W). Dry infected cocoa shoots ('dried brooms') were collected at the same site. Spore suspension was prepared using the protocol described by Frias and Purdy (1995). The isolate ESJOH-1 was maintained on potato–dextrose–agar (PDA) medium at room temperature.

Plant inoculation

Flushing apical meristems were inoculated by pipetting with 30 μL of a basidiospore suspension (10^5 spores/mL), followed by incubation in a humid chamber for 24 h. Seedlings from 'Catongo' were used as susceptibility controls to estimate inoculation efficiency. Shoot tips from inoculated and non-inoculated plants from genotypes 'CAB 214', 'CAB 208' and 'ICS 39' were collected at 24, 48, 72, 120 and 240 h after inoculation and frozen in liquid nitrogen (N_2) until RNA extraction. For each sampling period, eight plants were collected from each genotype plus a non-inoculated control. To evaluate inoculation efficiency, one inoculated plant and another non-inoculated from each sampling period were kept for evaluation of symptoms 30 days after inoculation. Disease incidence was estimated by the number of plants exhibiting any of the witches' broom symptoms (terminal and/or axillary brooms, length of terminal broom, swelling of stem, petiole and/or leaf pulvini, number of axillary brooms longer than 1 cm).

Plant and fungal RNA extraction

Total RNA from *T. cacao* shoot tips was extracted using the protocol described by Verica *et al.* (2004). Briefly, 2 g of tissue was ground in liquid N_2 and extracted with 15 mL buffer (2% CTAB, 2% PVP, 100 mM Tris pH 8.0, 5 mM EDTA, 2 M NaCl, 2% β-mercaptoethanol). After mixing, 15 mL of chloroform was added, mixed and centrifuged for 20 min at 8000*g* at 4 °C. Total RNA was precipitated from the supernatant by adding 8 M LiCl solution and incubating at 4 °C for 12 h, followed by centrifugation at 8000 ***g*** for 30 min. The pellet was resuspended and purified using RNAeasy (Quiagen, Hilden, Germany). Total RNA was extracted from 15-day-old cultures of *C. perniciosa* grown on potato dextrose broth using Trizol (Invitrogen, Carlsbad, CA). RNA quality and concentration were evaluated by 1.2% agarose gel electrophoresis and by spectrophotometry.

Construction of suppression subtractive libraries (SSH) from *T. cacao*

RNA samples from inoculated shoot tips from genotypes 'ICS 39' and 'CAB 214' were used to prepare two SSH libraries, using the PCR selected cDNA subtraction kit (Clontech, Mount View, CA). The cDNA was synthesized using 1 μg of a mixture of total RNA derived from the five collection periods for each genotype using the SuperSMART cDNA synthesis kit (Clontech). The approach adopted was cross-subtraction in both directions for inoculated genotypes. For the first SSH library, cDNA from the susceptible genotype ('ICS 39') was used as tester and from the resistant ('CAB 214') as driver, while for the second library the cDNA from 'CAB 214' was the tester and from 'ICS 39' was the driver. Amplification products were ligated to pGEM-T Easy (Promega, WI), transformed into electrocompetent DH10B *E. coli* cells, and plated into selective media. White colonies were isolated and individualized in microtitre plates (96-well) containing selective medium with 8% glycerol and stored at –80 °C. Amplification controls were conducted at the various stages in the construction of the libraries using the gene-specific primers, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene for *T. cacao* (designed based on TC385 at TIGR) and actin for *C. perniciosa* (Gesteira *et al.*, 2003).

Sequencing

A sample of clones from each plate was checked for the presence of inserts by amplification using M13 primers. Colony amplification products were directly sequenced. One microlitre from bacterial suspension in glycerol was diluted to 11 μL in sterile water, and denatured at 95 °C for 5 min. Amplification reactions contained 11 μL of plasmid solution, 50 mM KCl, 20 mM Tris-HCl pH 8.4, 1.5 mM $MgCl_2$, 100 μM of each dNTP, 0.2 μM of each M13 primer and 1.5 U *Taq* DNA polymerase (Invitrogen do Brazil, SP, Brazil) in a final volume of 19 μL. For sequencing, 1 μL of previously amplified product was used, employing either a T7 or an SP6 primer. Sequencing was performed using standard protocols of the DYEnamic ET Terminator Cycle Sequencing kit (Amersham Biosciences) in an ABI-3100 Genetic Analyser (Applied Biosystems, Foster City, CA).

Sequence analyses and annotation

Sequencing chromatograms were checked for quality, vector contamination by Phred/Phrap/Crossmatch (Ewing and Green, 1998; Ewing *et al.*, 1998) and clustered by CAP3 (Huang and Madan, 1999). The resulting sequences were compared with the *Arabidopsis thaliana* database (http://www.arabidopsis.org/) and with the non-redundant GenBank and dbEST databases from NCBI using BlastX. Sequences were also analysed against the tentative consensus from the *T. cacao* gene index at TIGR (http://www.tigr.org/) and compared with the database of the *C. perniciosa* genomic sequences (http://www.lge.ibi.unicamp.br/vassoura/; kindly authorized by Professor Gonçalo A. G. Pereira, UNICAMP, Brazil). Sequences were considered as 'no hit' when *e*-values were $> 10^{-3}$. To categorize sequences with positive significant matches, the Arabidopsis homologue gene codes were used to classify matching *T. cacao* sequences according to the ontology classification using The Arabidopsis Information Resource database (TAIR; www.arabidopsis.org). For those genes without identified homologues in the Arabidopsis genome, putative ontology was identified by a 'key-word' search of the database (http://www.geneontology.org). The sequence data described in this paper have been submitted to GenBank under accession numbers EH057645 to EH057831.

Validation of libraries by RT-qPCR

Primers were designed (see supplementary Table S2) based on sequences available at either SSH libraries and *T. cacao* sequences from the GenBank or TIGR Gene Index. The *T. cacao* gene for actin (TC46 from the *T. cacao* gene index at TIGR) was chosen as internal reference for quantitative expression analyses. Primers were designed using Primer3 and evaluated with NetPrimer for stability and mispairing.

Reverse transcription (RT) of a previously DNAse I-treated RNA aliquot from the same samples used to prepare both SSH libraries was conducted using SuperScript First Strand Synthesis System for RT-PCR (Invitrogen) with 18mer oligo-dT. Quantitative real-time amplifications from reversed-transcribed samples (RT-qPCR) of sample RNAs were conducted in 10-μL reactions containing 1 μL cDNA 1 : 10 (v/v) dilution, 0.5 μM of each transcript-specific primer (supplementary Table S2) and 5 μL platinum SYBR-green qPCR SuperMix-UDG 2X (Invitrogen). Amplification was performed in a RotorGene 3000 thermocycler (Corbett Life Science, Australia) in triplicates, with initial incubation at 50 °C for 2 min, 95 °C for 2 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 30 s, with fluorescence detection at the end of extension cycles. After cycling, melting curves for each amplicon were determined between 72 and 95 °C. Experiments included a negative control (no template DNA), and amplification efficiency of each primer pair was determined by a standard curve derived from three serial dilutions of cDNA (10^{-1}, 10^{-2}, 10^{-3}). Real-time data acquisition was effected with RotorGene Real-Time Analysis 6.0 (Corbett), which provided the values of cycle threshold (C_T) and PCR efficiency (E). The average C_T from the triplicates was used to determine differences in expression using the internal reference (actin) based on $2^{-(\Delta\Delta Ct)}$ (Livak and Schmittgen, 2001).

Evaluation of gene expression was initially conducted for non-inoculated samples from both genotypes ('ICS 39' and 'CAB 214') at the various sampling times (48, 72, 120, 240 h), contrasting both genotypes. Changes in gene expression in inoculated samples were compared with non-inoculated samples for the same genotype for each sampling period. This same analysis was conducted for samples from another resistant genotype ('CAB 208') using as reference a pool of cDNA from non-inoculated plants at all sampling periods. As the experiments did not contain biological replications, significant differences were estimated based on threshold as limit proposed by Benitez *et al.* (2005), with relative differences between 0.24 and 2.8 considered to be non-significant.

Detection of *C. perniciosa* transcripts in inoculated samples

The presence of *C. perniciosa* mycelia in *T. cacao* tissues was confirmed by RT-PCR. Specific primers for transcripts of the *C. perniciosa* ribosomal protein L35 gene (*RpL35*) (5′-ACTTCGGGTGCAAAAGATTG-3′; 5′-TGGTCCTTCTTCGTCTGCTT-3′) were designed based on sequences deposited at the *C. perniciosa* genomic database (http://www.lge.ibi.unicamp.br/vassoura/). cDNA samples from infected cocoa tissue collected at 24, 48, 72, 120 and 240 h were analysed by RT-qPCR for *RpL35*, using cDNA (synthesized as above) derived from pure culture of *C. perniciosa* as positive control. Transcript detection was based on specific melting curves of amplified products and by 1.5% ethidium bromide gel electrophoresis.

ACKNOWLEDGEMENTS

We wish to thank FAPESP for financial support (03/11483-0), CNPq for fellowships and CEPLAC for technical support. Technical assistance by Edivaldo Pimentel was greatly appreciated. We would like to thank the reviewers for their critical comments.

REFERENCES

Albuquerque, P.S.B. and Figueira, A. (2004) Genetic mapping and identification of genomic regions associated with witches' broom resistance, derived from alternative sources from the Brazilian Amazon (CAB genotypes). *INGENIC Newsletter*, **9**, 39–42.

Andebrhan, T., Figueira A., Yamada. M.M., Cascardo J. and Furtek D.B. (1999) Molecular fingerprinting suggests two primary outbreaks of witches' broom disease (*Crinipellis perniciosa*) of *Theobroma cacao* in Bahia, Brazil. *Eur. J. Plant Pathol.* **105**, 167–175.

Aneja, M. and Gianfagna, T. (2001) Induction and accumulation of caffeine in young, actively growing leaves of cocoa (*Theobroma cacao* L.) by wounding or infection with *Crinipellis perniciosa*. *Physiol. Mol. Plant Pathol.* **59**, 13–16.

Bailey, B.A., Bae, H., Strem, M.D., Mayolo, G.A., Guiltinan, M.J., Verica, J.A., Maximova, S.N. and Bowers, J.H. (2005) Developmental expression of stress response genes in *Theobroma cacao* leaves and their response to Nep1 treatment and a compatible infection by *Phytophthora megakarya*. *Plant Physiol. Biochem.* **43**, 611–622.

Basse, C.W. (2005) Dissecting defense-related and developmental transcriptional responses of maize during *Ustilago maydis* infection and subsequent tumor formation. *Plant Physiol.* **138**, 1774–1784.

Benitez, Y., Botella, M.A., Trapero, A., Alsalimiya, M., Caballero, J.L., Dorado, G. and Blanco, J.M. (2005) Molecular analysis of the interaction between *Olea europaea* and the biotrophic fungus *Spilocaea oleagina*. *Mol. Plant Pathol.* **6**, 425–438.

Bertini, L., Cascone, A., Tucci, M., D'Amore, R., Di Berardino, I., Buonocore, V., Caporale, C. and Caruso, C. (2006) Molecular and functional analysis of new members of the wheat PR4 gene family. *Biol. Chem.* **387**, 1101–1111.

Bowers, J.H., Bailey, B.A., Hebbar, P.K., Sanogo, S. and Lumsden, R.D. (2001) The impact of plant diseases on world chocolate production. *Plant Health Prog.* doi: 10.1094/PHP-2001-0709-01-RV.

Brownlee, H.E., Hedger, J.N. and Scott, I.M. (1992) Effect of a range of procyanidins on the cocoa pathogens *Crinipellis perniciosa*. *Physiol. Mol. Plant Pathol.* **40**, 227–232.

Chapple, C. (1998) Molecular-genetic analysis of plant cytochrome p450-dependent monooxygenases. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 311–343.

Chittoor, J.M., Leach, J.E. and White, F.F. (1997) Differential induction of a peroxidase gene family during infection of rice by *Xanthomonas oryzae* pv. *oryzae*. *Mol. Plant–Microbe Interact.* **10**, 861–871.

Collinge, D.B., Kragh, K.M., Mikkelsen, J.D., Nielsen, K.K., Rasmussen, U. and Vad, K. (1993) Plant chitinases. *Plant J.* **3**, 31–40.

Degenhardt, J., Al-Masri, A.N., Kürkcüoglu, S., Szankowski, I. and Gau, A.E. (2004) Characterization by suppression subtractive hybridization of transcripts that are differentially expressed in leaves of apple scab-resistant and susceptible cultivars of *Malus domestica*. *Mol. Gen. Genomics*, **273**, 3226–3335.

Delgado, J.C. and Cook, A.A. (1976) Nuclear condition of the basidia basidiospore and mycelial of *Marasmius perniciosus*. *Can. J. Bot.* **54**, 66–72.

Diatchenko, L., Lau, Y.F.C., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, L., Lukyanov, L., Gurskaya, N., Sverdlov, E.D. and Sieber, P. (1996) Suppression subtractive hybridization: a method for generating differentially or tissue-specific cDNA probes and libraries. *Proc. Natl Acad. Sci. USA*, **93**, 6025–6030.

Ewing, B. and Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M. and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185.

Faleiro, F.G., Queiroz, V.T., Lopes, U.V., Guimarães, C.T., Pires, J.L., Yamada, M.M., Araújo, I.S., Pereira, M.G., Schnell, R., de Souza, G.A., Ferreira, C.F., Barros, E.G. and Moreira, M.A. (2006) Mapping QTLs for witches' broom (*Crinipellis perniciosa*) resistance in cacao (*Theobroma cacao* L.). *Euphytica*, **149**, 227–235.

Fernandez, D., Santos, P., Agostini, C., Bom, M.C., Petitot, A.S., Silva, M.C., Guerra-Guimarães, L., Ribeiro, A., Argout, X. and Nicole, M. (2004) Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). *Mol. Plant Pathol.* **5**, 527–536.

Frias, G.A. and Purdy, L.H. (1995) An inoculation method for evaluating resistance of cocoa to *Crinipellis perniciosa*. *Plant Dis.* **79**, 787–791.

Frias, G.A., Purdy, L.H. and Schmidt, R.A. (1991) Infection biology of *Crinipellis perniciosa* on vegetative flushes of cacao. *Plant Dis.* **75**, 552–556.

Gesteira, A.S., Micheli, F., Ferreira, C.F. and Cascardo, J.M.C. (2003) Isolation and purification of functional total RNA from different organs of cacao tree during its interaction with the pathogen *Crinipellis* perniciosa. *Biotechniques*, **35**, 494–496.

Guerreiro Filho, O. and Mazzafera, P. (2000) Caffeine does not protect coffee against the leaf miner *Perileucoptera coffeella*. *J. Chem. Ecol.* **26**, 1447–1464.

Guerreiro Filho, O. and Mazzafera, P. (2003) Caffeine and resistance of coffee to the berry borer *Hypothenemus hampei* (coleoptera: scolytidae). *J. Agric. Food Chem.* **51**, 6987–6991.

Guilleroux. M. and Osbourn, A. (2004) Gene expression during infection of wheat roots by the 'take-all' fungus *Gaeumannomyces graminis*. *Mol. Plant Pathol.* **5**, 203–216.

Hohnjec, N., Küster, H., Albus, U., Frosch, S.C., Becker, J.D., Pühler, A., Perlick, A.M. and Frühling, M. (2000) The broad bean nodulin VfENOD18 is a member of a novel family of plant proteins with homologies to the bacterial MJ0577 superfamily. *Mol. Gen. Genet.* **264**, 241–250.

Holter, N.S., Mitra, M., Maritan, M., Cieplak, M., Bnavar, J.R. and Fedoroff, N.V. (2000) Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *Proc. Natl Acad. Sci. USA*, **97**, 8409–8414.

Huang, X. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* **9**, 868–877.

Jones, P.G., Allaway, D., Gilmour, D.M., Harris, C., Rankin, D., Retzel, E.R. and Jones, C.A. (2002) Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. *Planta*, **216**, 255–264.

Kauffmann, S., Legrand, M., Geoffroy, P. and Fritig, B. (1987) Biological function of 'pathogenesis-related' proteins: four PR proteins of tobacco have 1,3-β-glucanase activity. *EMBO J.* **6**, 3209–3212.

Kawaoka, A., Matsunaga, E., Endo, S., Kondo, S., Yoshida, K., Shinmyo, A. and Ebinuma, H. (2003) Ectopic expression of a horseradish peroxidase enhances growth rate and increases oxidative stress resistance in hybrid aspen. *Plant Physiol.* **132**, 1177–1185.

Kazan, K., Schenk, P.M., Wilson, I. and Manners, J.M. (2001) DNA microarrays: new tools in the analysis of plant defence response. *Mol. Plant Pathol.* **2**, 177–185.

Kerk, D., Bulgrien, J., Smith, D.W. and Gribskov, M. (2003) Arabidopsis proteins containing similarity to the universal stress protein domain of bacteria. *Plant Physiol.* **131**, 1209–1219.

Laker, H.A., Sreenivasan, T.N. and Raj-Kumar, D. (1988) Present status of witches' broom disease of cocoa in Trinidad. *Trop. Pest Manage.* **34**, 318–323.

Lee, S.C., Kim, S.H., An, S.H., Yi, S.Y. and Hwang, B.K. (2006) Identification and functional expression of the pepper pathogen induced gene, CAPIP2, involved in disease resistance and drought and salt stress tolerance. *Plant Mol. Biol.* **62**, 151–164.

Livak, K.L. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, **25**, 402–408.

Logemann, E., Wu, S.C., Schroder, J., Schmelzer, E., Somssich, I.E. and Hahlbrock, K. (1995) Gene activation by UV light, fungal elicitor or fungal infection in *Petroselinum crispum* is correlated with repression of cell cycle-related genes. *Plant J.* **8**, 865–876.

Lu, G., Jantasuriyarat, C., Zhou, B. and Wang, G.L. (2004) Isolation and characterization of novel defense response genes involved in compatible and incompatible interactions between rice and *Magnaporthe grisea*. *Theor. Appl. Genet.* **108**, 525–534.

Ng, M. and Yanofsky, M.F. (2001) Activation of the arabidopsis B class homeotic genes by *APETALA1*. *Plant Cell* **13**, 739–753.

Podila, G.K., Zheng, J., Balasubramanian, S., Sundaram, S., Hirematths, S., Brand, J.H. and Hymes, M.J. (2002) Fungal gene expression in early symbiotic interactions between *Laccaria bicolor* and red pine. *Plant Soil*, **244**, 117–128.

Purdy, L.H. and Schimdt, R.A. (1996) Status of cacao witches broom: biology, epidemiology, and management. *Annu. Rev. Phytopathol.* **34**, 573–594.

Quiroga, M., Guerrero, C., Botella, M.A., Barceló, A., Amaya, I., Medina, M.I., Alonso, F.J., de Forchetti, S.M., Tigier, H. and Valpuesta, V. (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. *Plant Physiol.* **122**, 1119–1127.

Ross, J.R., Nam, K.H., D'Auria, J.C. and Pichersky, E. (1999) *S*-adenosyl-1-methionine: salicylic acid carboxyl methyltransferase, an enzyme involved in floral scent production and plant defense, represents a new class of plant methyltransferases. *Arch. Biochem. Biophys.* **367**, 9–16.

Sasaki, K., Iwai, T., Hiraga, S., Kuroda, K., Seo, S., Mitsuhara, I., Miyasaka, A., Iwano, M., Ito, H., Matsui, H. and Ohashi, Y. (2004) Ten rice peroxidases redundantly respond to multiple stresses including infection with rice blast fungus. *Plant Cell Physiol.* **45**, 1442–1452.

Sauter, M., Rzewuski, G., Marwedel, T. and Lorbiecke, R. (2002) The novel ethylene-regulated gene OsUsp1 from rice encodes a member of a plant protein family related to prokaryotic universal stress proteins. *J. Exp. Bot.* **53**, 2325–2331.

Scarpari, L.M., Meinhardt, L.W., Mazzafera, P., Pomella, A.W.V., Schiavinato, M.A., Cascardo, J.C.M. and Pereira, G.A.G. (2005) Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa*. *J. Exp. Bot.* **56**, 865–877.

Seo, H.S., Song, J.T., Cheong, J., Lee, Y., Lee, Y., Hwang, I., Lee, L.S. and Choi, Y.D. (2001) Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. *Proc. Natl Acad. Sci. USA*, **98**, 4788–4793.

Sereno, M.L., Albuquerque, P.S.B., Vencovsky, R. and Figueira, A. (2006) Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conserv. Genet.* **7**, 13–24.

Silva, S.D.V.M., Luz, E.D.M.N. and Matsuoka, K. (1998) Seleção de variáveis na avaliação de progênies de cacaueiro (*Theobroma cacao*) quanto a resistência a *Crinipellis perniciosa*. *Agrotrópica*, **10**, 87–94.

Silva, S.D.V.M. and Matsuoka, K. (1999) Histologia da interação *Crinipellis perniciosa* em cacaueiros suscetível e resistente à vassoura-de-bruxa. *Fitopatol. Brasileira*, **24**, 54–59.

Somssich, I.E. and Hahlbrock, K. (1998) Pathogen defence in plants—a paradigm of biological complexity. *Trends Plant Sci.* **3**, 86–90.

Takemoto, D., Doke, N. and Kawakita, K. (2001) Characterization of elicitor-inducible tobacco genes isolated by differential hybridization. *J. Gen. Plant Pathol.* **67**, 89–96.

Tao, Y., Xie, Z., Chen, W., Glazebrook, J., Chang, H.-S., Han, B., Zhu, T., Zou, G. and Katagiri, F. (2003) Quantitative nature of Arabidopsis responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell*, **15**, 317–330.

Thara, V.K., Fellers, J.P. and Zhou, J.M. (2003) In planta induced genes of *Puccinia triticina*. *Mol. Plant Pathol.* **4**, 51–56.

Verica, J.A., Maximova, S.N., Strem, M.D., Carlson, J.E., Bailey, B.A. and Guiltinan, M.J. (2004) Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. *Plant Cell Rep.* **23**, 404–413.

Voiblet, C., Duplessis, S., Encelot, N. and Martin, F. (2001) Identification of symbiosis-regulated genes in *Eucalyptus globulus-Pisolithus tinctorius* ectomycorrhiza by differential hybridization of arrayed cDNAs. *Plant J.* **25**, 181–191.

Wheeler, B.E.J. (1985) The growth of *Crinipellis perniciosa* in living and dead cocoa tissue. In: *Developmental Biology of Higher Fungi* (Moore, D., Casselton, L.A., Wood, D.A. and Frankland, J.C., eds), pp. 103–116. Cambridge: Cambridge University Press.

Winkel-Shirley, B. (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* **126**, 485–493.

Zuo, K., Wang, J., Wu, W., Chai, Y., Sun, X. and Tang, K. (2005) Identification and characterization of differentially expressed ESTs of *Gossypium barbadense* infected by *Verticillium dahliae* with suppression subtractive hybridization. *Mol. Biol.* **39**, 191–199.

SUPPLEMENTARY MATERIAL

The authors have provided the following supplementary material which can be accessed alongside the article on http://www.blackwell-synergy.com

Table S1. Annotation and *E*-values from unique sequences from the SSH libraries enriched for inoculated susceptible (ICS 39) sequences or inoculated resistant (CAB 214) based on BLASTX analyses against the TAIR, NCBI, or TIGR database. Sequences were grouped according to functional category.

Table S2. Genes selected to be validated by quantitative reverse transcription amplification, with respective putative annotation; NCBI GenBank or tentative consensus (TC) identification from TIGR (www.tigr.org); and primer sequences.