CSIRO PUBLISHING

Functional Plant Biology, 2013, **40**, 1029–1047
http://dx.doi.org/10.1071/FP12296

Transcriptome analyses and virus induced gene silencing identify genes in the *Rpp4*-mediated Asian soybean rust resistance pathway

Aguida M. A. P. Morales[A], *Jamie A. O'Rourke*[B], *Martijn van de Mortel*[C], *Katherine T. Scheider*[D], *Timothy J. Bancroft*[E], *Aluízio Borém*[A], *Rex T. Nelson*[F], *Dan Nettleton*[E], *Thomas J. Baum*[C], *Randy C. Shoemaker*[F], *Reid D. Frederick*[D], *Ricardo V. Abdelnoor*[G], *Kerry F. Pedley*[D], *Steven A. Whitham*[C] *and Michelle A. Graham*[F,H,I]

[A]Universidade Federal de Viçosa, Departamento de Fitotecnia, 36.570-000, Viçosa, MG, Brazil.
[B]USDA-Agricultural Research Service, Plant Science Research Unit, Saint Paul, MN 55108, USA.
[C]Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA 50014, USA.
[D]USDA- Agricultural Research Service, Foreign Disease-Weed Science Research Unit, Fort Detrick, MA 21702, USA.
[E]Department of Statistics, Iowa State University, Ames, IA 50014, USA.
[F]USDA-Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, Ames, IA 50014, USA.
[G]Laboratório de Biotecnologia Vegetal e Bioinformática, Embrapa Soja, Rod. Carlos João Strass, 86001-970, Londrina – PR, Brazil.
[H]Department of Agronomy, Iowa State University, Ames, IA 50014, USA.
[I]Corresponding author. Email: michelle.graham@ars.usda.gov

Abstract. *Rpp4* (*Resistance to Phakopsora pachyrhizi 4*) confers resistance to *Phakopsora pachyrhizi* Sydow, the causal agent of Asian soybean rust (ASR). By combining expression profiling and virus induced gene silencing (VIGS), we are developing a genetic framework for *Rpp4*-mediated resistance. We measured gene expression in mock-inoculated and *P. pachyrhizi*-infected leaves of resistant soybean accession PI459025B (*Rpp4*) and the susceptible cultivar (Williams 82) across a 12-day time course. Unexpectedly, two biphasic responses were identified. In the incompatible reaction, genes induced at 12 h after infection (hai) were not differentially expressed at 24 hai, but were induced at 72 hai. In contrast, genes repressed at 12 hai were not differentially expressed from 24 to 144 hai, but were repressed 216 hai and later. To differentiate between basal and resistance-gene (*R*-gene) mediated defence responses, we compared gene expression in *Rpp4*-silenced and empty vector-treated PI459025B plants 14 days after infection (dai) with *P. pachyrhizi*. This identified genes, including transcription factors, whose differential expression is dependent upon *Rpp4*. To identify differentially expressed genes conserved across multiple *P. pachyrhizi* resistance pathways, *Rpp4* expression datasets were compared with microarray data previously generated for *Rpp2* and *Rpp3*-mediated defence responses. Fourteen transcription factors common to all resistant and susceptible responses were identified, as well as fourteen transcription factors unique to *R*-gene-mediated resistance responses. These genes are targets for future *P. pachyrhizi* resistance research.

Additional keywords: defence, genomics, resistant genes, *Rpp4*, soybean.

Received 5 October 2012, accepted 12 January 2013, published online 15 July 2013

Introduction

Asian soybean rust (ASR) is caused by the obligate fungus *Phakopsora pachyrhizi* Sydow. ASR is considered polycyclic, since *P. pachyrhizi* is able to complete several generations in a single life cycle of the host (Yorinori *et al*. 2005). Temperatures that favour the growth and development of soybean also favour the development of rust. The disease destroys leaf tissue, resulting in reduced photosynthetic activity, premature defoliation and shortened life cycle. The cumulative effect of rust on soybean production translates into lower seed weight and reduces the number of pods and seeds (Sinclair 1989). Currently, *P. pachyrhizi* is one of the most important economic threats for soybean growers in South America. In Brazil, a 2-year field trial demonstrated that rust was responsible for 37–67% of soybean seed yield losses (Kumudini *et al*. 2008).

Six genes for resistance to *P. pachyrhizi* (*Rpp*) have been identified in soybean: *Rpp1*, *Rpp2*, *Rpp3*, *Rpp4*, *Rpp5* and

Rpp6 (Bromfield and Hartwig 1980; McLean and Byth 1980; Hartwig and Bromfield 1983; Hartwig 1986; Silva *et al*. 2008; Garcia *et al*. 2008; Li *et al*. 2012). Each of these genes has been genetically mapped in the soybean genome (Hyten *et al*. 2007, 2009; Monteros *et al*. 2007; Garcia *et al*. 2008; Silva *et al*. 2008; Chakraborty *et al*. 2009; Kim *et al*. 2012). Further, comparisons made between molecular markers used to map *Rpp1-Rpp4* and the soybean genome sequence (Schmutz *et al*. 2010) identified clusters of candidate resistance gene (*R*-gene) homologues (Graham 2012). However, identifying the actual *R*-genes remains difficult. Among the six *Rpp* genes, *Rpp4* is the best characterised and has been the most stable when challenged against isolates from different parts of the world (Yamaoka *et al*. 2002; Bonde *et al*. 2006). Sequencing of the 2 cM region in the susceptible cultivar 'Williams82' (Wm82) revealed a cluster of three coiled-coil nucleotide-binding site leucine-rich repeats (CC-NBS-LRR) *R*-genes with similarity to the lettuce RGC2 family of NBS-LRR resistance genes (Meyer *et al*. 2009). Expression analyses of these genes in the resistant accession PI459025B and susceptible cultivar Wm82 revealed large differences in gene number and expression. Virus induced gene silencing (VIGS) performed using constructs developed from the Wm82 *Rpp4* candidate genes silenced *Rpp4*-mediated resistance in PI459025B, indicating *Rpp4* is a member of the same gene cluster.

Several groups have used transcriptomic approaches to characterise compatible and incompatible responses to *P. pachyrhizi* (Panthee *et al*. 2007; van de Mortel *et al*. 2007; Choi *et al*. 2008; Panthee *et al*. 2009; Tremblay *et al*. 2010, 2012). Perhaps the most comprehensive transcriptomic studies thus far have been on *Rpp2*-mediated defence. van de Mortel *et al*. (2007) examined a 7-day time course of *P. pachyrhizi* infection in a resistant (*Rpp2*-mediated) and susceptible genotype. A biphasic response to *P. pachyrhizi* was observed in both genotypes. At 12 h after inoculation (hai) with *P. pachyrhizi*, basal defence was induced in both genotypes. However, at 24 hai, defence gene expression returned to mock-inoculated levels. At 72 hai, a second round of defence gene expression occurred in the resistant (incompatible) reaction, likely attributable to *Rpp2*-mediated signalling. Although this secondary defence response was also detected in the susceptible (compatible) reaction, it occurred later (96 hai) and never to the ame magnitude observed in the resistant reaction. Although greater levels of defence-related gene induction were observed in the resistant reaction, greater numbers of differentially expressed genes were observed in the susceptible reaction.

This work was followed by Pandey *et al*. (2011), who identified genes required in the *Rpp2*-mediated signalling pathway. The authors used the work by van de Mortel *et al*. (2007) and soybean orthologs of known defence genes to identify 140 genes potentially involved in *Rpp2*-mediated defence. VIGS of these genes followed by *P. pachyrhizi* infection was used to identify 11 genes essential for *Rpp2*-mediated resistance. These included four soybean orthologs of known defence genes (*GmEDS1*, *GmNRP1*, *GmPAD4* and *GmPAL1*), five predicted transcription factors (*GmWRKY36*, *GmWRKY40*, *GmWRKY45*, *GmDBTF* and *GmMYB84*), an O-methyl transferase (*GmO-MT*) and a cytochrome P450 (*GmCYP83E12*). Combining these results with data from other plant–pathogen systems allowed for the characterisation of the *Rpp2* signalling cascade, even though *Rpp2* has yet to be cloned.

Recently, our group used microarray analyses and microscopy to characterise resistance and susceptibility to *P. pachyrhizi* in PI462312 (*Rpp3*, Schneider *et al*. 2011). As in the case of the *Rpp2* experiments (van de Mortel *et al*. 2007), *P. pachyrhizi* elicits a biphasic response characterised by a burst of differential gene expression in the first 12 hai, correlated with fungal appressorium formation and penetration of epidermal cells (Schneider *et al*. 2011). A quiescent period occurs from 24 to 48 hai, in which *P. pachyrhizi* continued to develop but did not elicit strong host responses, followed by a second phase of intense defence gene expression. The second burst of gene expression followed the onset of haustoria formation in both compatible and incompatible interactions. The temporal relationship between *P. pachyrhizi* growth and host responses provides an important context in which to view interacting gene networks that mediate the outcomes of the plant-pathogen interaction.

In this study, the power of genomic, transcriptomic, bioinformatic and VIGS approaches were used to identify genes in the *Rpp4*-mediated ASR resistance pathway. Microarray analysis of the resistant soybean accession PI459025B (*Rpp4*) and the susceptible soybean cultivar Wm82, inoculated or mock-inoculated with the Hawaii94–1 (Hw94–1) isolate of *P. pachyrhizi*, was conducted to establish a twelve-day time course of soybean gene expression. Genes induced or repressed by *P. pachyrhizi* infection in PI459025B have different biphasic expression patterns. To differentiate between basal and *R*-gene mediated defence responses, microarray analysis was used to compare gene expression in *Rpp4*-silenced and mock-silenced plants. In addition, we compared these data with the *Rpp2* and *Rpp3* microarray data developed by our group (van de Mortel *et al*. 2007; Schneider *et al*. 2011) to identify transcription factors associated with multiple *R*-gene mediated pathways.

Materials and methods

Inoculation and tissue collection from P. pachyrhizi *resistant and susceptible cultivars*

Phakopsora pachyrhizi Sydow inoculations were performed at the USA Department of Agriculture-Agricultural Research Service Foreign Disease-Weed Science Research Unit (FDWSRU) Biological Safety Level-3 Plant Pathogen Containment Facility at Fort Detrick, Maryland (Melching *et al*. 1983). Urediniospores of the *P. pachyrhizi* isolate HW94–1 were removed from liquid nitrogen, heat shocked at 40°C for 5 min, then hydrated at 100% RH at room temperature for 16 h. Spores were suspended in sterile distilled water containing 0.01% (v/v) Tween 20, filtered through a 53-μm pore-size screen, and adjusted to a concentration of 5×10^5 spores mL^{-1} by means of a hemacytometer. Three plants per pot were inoculated at the V3 growth stage by spraying ~10 mL of urediniospore solution on the adaxial surface of the leaves of the soybean accession PI459025B or cultivar

Wm82. The accession PI459025B (*Rpp4*) forms red-brown lesions upon infection with *P. pachyrhizi* isolate HW94–1, indicating a resistance response, whereas the susceptible cultivar Wm82 forms tan lesions with sporulating uredinia (Bonde *et al*. 2006). Mock treatments used the same solution minus *P. pachyrhizi* urediniospores. Following inoculation, plants were placed in dew chambers for 16 h at 20°C, and were then moved to the greenhouse maintained at 20°C. Supplemental illumination was provided by 1000-W Metalarc lights (Sylvania, Danvers, MA, USA) spaced 0.6 m apart above the bench.

Three leaflets of the third trifoliolate leaves of three plants from each cultivar and treatment were collected at 12, 24, 72, 144, 216, and 288 hai. The leaves were frozen in liquid nitrogen and stored at −80°C. RNA was isolated from 200 mg of ground leaf tissue using 1 mL of Tri Reagent (Molecular Research Center, Cincinnati, OH, USA). An overnight precipitation in 2 M (final concentration) lithium chloride at −20°C was used to further purify samples (Ausubel *et al*. 1994) followed by RNeasy column purification (Qiagen, Valencia, CA, USA) and elution in 30 µL of diethyl pyrocarbonate-treated water. For simplicity, this experiment will be referred to as the *Rpp4* time course dataset.

Inoculation and tissue collection from Rpp4*-silenced and mock-silenced PI459025B plants infected with* P. pachyrhizi

Previously, our group used VIGS to silence *Rpp4* candidate genes in PI459025B (Meyer *et al*. 2009). The plants used in our analyses here are the same plants used in the study by Meyer *et al*. (2009). In brief, a portion of the LRR domain of the *Rpp4* candidate genes from Wm82 was cloned into RNA2 of the *Bean pod mottle virus* (BPMV) VIGS vector (BPMV-LRR) (Meyer *et al*. 2009). Co-bombardment of BPMV RNAs 1 and 2 on Wm82 leaves was used to generate inoculum for further experimentation. After 3 weeks, BMPV-infected tissue was collected, lyophilised and shipped to the FDWSRU. The resistant soybean accession PI459025B was grown in a growth chamber at 20°C with 16 h light, and 2 weeks after germination, plants were rub-inoculated with VIGS constructs. Each construct was tested on six plants. Controls included non-treated plants, mock inoculated plants and plants treated with empty BPMV constructs. Three weeks later, plants were inoculated with *P. pachyrhizi* isolate LA04–1. This isolate also results in the formation of resistant red-brown lesions in PI459025B. This isolate was used for this experiment because it was derived from a single pustule. By using VIGS and LA04–1, we minimised background plant and pathogen genomic effects and identified genes that are regulated in an *Rpp4*-dependent manner. Two weeks later, plants were evaluated for resistance or susceptibility to *P. pachyrhizi* strain LA04–1. Three independent replicates of the experiment were performed. After the completion of each replicate experiment, leaves were collected from three BPMV-LRR VIGS plants and three empty-vector BPMV plants, all infected with *P. pachyrhizi*. This provided three biological replicates to use for microarray analyses. Leaves were flash frozen in liquid nitrogen and stored at −80°C. RNA was isolated from ground leaf tissue using the Plant RNeasy kit (Qiagen) and subsequently treated with DNase. For simplicity, data generated from these samples will be referred to as the *Rpp4* VIGS dataset.

Microarray labelling, hybridisation, and scanning

Labelling, hybridisation, and scanning of all tissue samples were performed at the Iowa State University GeneChip Facility. Labelled target copy RNA (cRNA) was synthesised from 5 µg of total RNA using the GeneChip One-Cycle Target Labelling and Control Reagents kit (Affymetrix, Santa Clara, CA, USA) according to manufacturer's instructions. Fragment cRNA (10 µg) were hybridised to GeneChip Soybean Genome Array (Affymetrix) according to manufacturer's instructions. cRNA quality was verified on an Agilent 2100 BioAnalyzer equipped with an RNA Nano LabChip. Microarrays were scanned with a GCS3000 7G scanner (Affymetrix).

Statistical analysis of microarray data

Our experimental samples described above differed greatly in complexity. In the time-course dataset we compared gene expression in resistant and susceptible cultivars, infected and mock-infected with *P. pachyrhizi* over six time points. In the VIGS dataset, we compared gene expression in *Rpp4*-silenced and mock-silenced plants at a single time point and genotype. Given the differences in these experiments, different methods were used to identify significantly differentially expressed probe sets.

In the *Rpp4* time course dataset, differentially expressed probe sets were identified using the methods described by Schneider *et al*. (2011). In short, linear model analysis of normalised log-scale expression measures was performed separately for each gene, using ver. 9.2 of SAS (SAS Institute, Cary, NC, USA). Each linear model included fixed effects for replicates, times, infection types and interactions between times and infection types. To determine whether there was significant evidence of an expression difference between infection types when averaged over time, whether there was significant evidence that the pattern of expression over time differed with infection type, and whether infection types differed significantly within individual time points, *F*-tests were conducted for infection type main effects, infection type by time interactions, and all possible comparisons of infection types at each time point. Together these tests identified genes whose expression differed in some manner (either in level, pattern over time, or both) between infection types within each genotype. A *q*-value was computed for each *F*-test *P*-value using the method described by Storey and Tibshirani (2003). The *q*-values were used to produce lists of differentially expressed genes with estimated false discovery rates (FDR) of 0.01% (see Tables S1 and S2, available as Supplementary Material to this paper).

For identifying differentially expressed probe sets in the VIGS dataset, raw expression values from the CEL files generated during array processing were read into R (R Development Core Team 2006). The data was analysed in the 'affy' background, corrected using the RMA function, normalised using the invariant set and summarised using the median polish command. A linear model analysis was

conducted for each gene using the limma package (Smyth 2005). The empirical Bayesian approach described by Smyth (2004) was used to moderate probe-wise sample variances. The *P*-values from tests of interest were converted to *q*-values to allow for the approximate control of FDR. Genes were declared to be differentially expressed when the absolute value of the estimated log 2-fold change exceeded 1 and the *q*-value was less <0.05 (Table S3).

For heat map generation, Cluster 3.0 was used for hierarchical clustering using the Pearson correlation (de Hoon *et al*. 2004). Java TreeView (Saldanha 2004) was used for heatmap visualisations.

Annotation of differentially expressed probe sets

The differentially expressed probe sets were annotated using the SoyBase Affymetrix GeneChip Soybean Genome Array Annotation page (ver. 3) as described by Schneider *et al*. (2011) (http://soybase.org/AffyChip, accessed October 2012; see Tables S1, S2 and S3). Annotation data included the corresponding Glyma1 identifier from the soybean genome sequence (Schmutz *et al*. 2010), the top three hits to the Uniprot protein database (Apweiler *et al*. 2004), the top *Arabidopsis* hit and the corresponding gene ontology (GO; Ashburner *et al*. 2000) biological process and molecular function terms inferred from the top *Arabidopsis* hit. To identify over-represented GO terms, Fisher's exact test (Fisher 1966) with a Bonferroni (1935) correction was used to compare GO term enrichment within the differentially expressed genes and the entire Soybean genome array.

Differentially expressed probe sets were also annotated using the SoyDB (http://casp.rnet.missouri.edu/soydb, accessed February 2011; Wang *et al*. 2010) transcription factor database based on the corresponding soybean gene identifier. Fisher's exact test (Fisher 1966) with a Bonferroni correction (Bonferroni 1935) was used to identify over-represented transcription factor classes.

Identification of significantly over-represented transcription factor binding sites

Each of the differentially expressed genes identified was assigned to a cDNA from the whole soybean genome assembly (Schmutz *et al*. 2010) using BLASTN (Altschul *et al*. 1997; E < 10–30, per cent identity >95). Using the coordinates of the corresponding soybean gene, custom perl scripts were used to extract 1000 bases of promoter sequence from whole soybean genome assembly. Clover (Cis-eLement Over-representation, Frith *et al*. 2004), in combination with the TRANSFAC transcription factor database (Wingender *et al*. 1996), was used to identify over-represented transcription factor binding sites in the promoters of the differentially expressed genes. As a background control, the results were compared with 1000 bases of promoter sequence from all predicted soybean genes excluding transposable elements. In order to identify transcription factors associated with particular GO terms, the GO information was used. Clover and TRANSFAC were used to compare promoters from genes within a GO category to all promoters in the genome.

Results

Identification of differentially expressed probe sets from resistant and susceptible genotypes infected with P. pachyrhizi *over time*

Plants of the resistant soybean accession PI459025B (*Rpp4*) and susceptible cultivar Wm82 were inoculated with *P. pachyrhizi* isolate HW94–1. Leaf samples were collected at 12, 24, 72, 144, 216 and 288 hai. The abundance of soybean mRNA transcripts was assayed using the Gene ChipSoybean Genome Array (Affymetrix) and these results were deposited in the Plant Expression Database (Accession number GM37; Wise *et al*. 2007). Significantly differentially expressed probe sets were identified from the six different time points using a 0.01% FDR ($q \leq 0.0001$, Nettleton 2006). At the 0.01% FDR, probe sets were identified that had expression patterns differing over time in *P. pachyrhizi*-infected plants compared with mock-inoculated control plants. In the resistant soybean accession PI459025B, 5805 differentially expressed probe sets were identified (Table S1) whereas in the susceptible cultivar Wm82, 5304 differentially expressed probe sets were identified (Table S2).

Annotation of P. pachyrhizi*-responsive genes from resistant and susceptible genotypes*

Gene descriptions and functional annotations were downloaded from the SoyBase Affymetrix Soybean Genome array website (http://www.soybase.org/AffyChip, accessed October 2012). Annotation information included corresponding gene models from the soybean genome sequence (Schmutz *et al*. 2010). Of the 5805 differentially expressed probe sets identified in the resistant reaction, 631 (10.9%) had no match to predicted genes in the genome sequence ($E < 10^{-30}$ and percent identity greater than 95%) and 132 (0.6%) could not be unambiguously assigned to a single soybean gene call. The remaining probes were assigned to 4136 unique soybean genes (Table S1). Similarly, for the 5304 differentially expressed probes identified in the susceptible reaction, 542 (10.2%) had no match to predicted genes in the soybean genome and 123 (2.3%) could not be unambiguously assigned to a single gene. The remaining probes were assigned to 4038 predicted genes (Table S2). For reference purposes, 2.2% of probes on the entire chip could not be unambiguously assigned to single gene call and 16.3% of probes on the chip had no match to predicted genes in the soybean genome.

In order to understand biological processes associated with resistance and susceptibility, we used GO information to identify terms significantly over-represented in our datasets. Ten GO biological process terms were significantly over-represented only in the resistant reaction including; defence response, defence response to fungus, responses to oxidative stress, chitin, other organisms, and salicylic acid stimulus, regulation of transcription and defence responses, photosynthesis and oxidation-reduction process (Fig. 1). Four GO terms were over-represented in both susceptible and resistant reactions, including responses to wounding, karrikin and jasmonic acid stimulus, and photosynthetic electron transport. Five GO terms including translation, ribosome biogenesis, chloroplast organisation, photosynthesis light reaction and metabolic

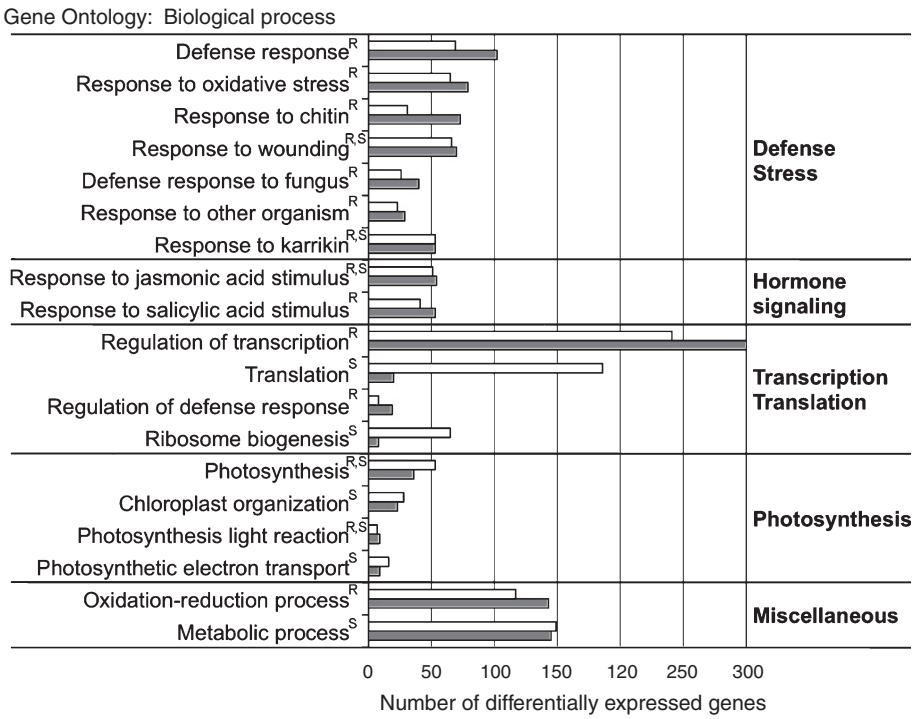

Fig. 1. Over-represented GO biological process classification of *Phakopsora pachyrhizi*- responsive genes as determined by Fisher's exact test (Fisher 1966). Differentially expressed genes (DEG) in resistant and susceptible reactions are represented by grey and white bars respectively. Since a DEG could be associated with multiple GO processes, it could be represented multiple times. Individual GO categories have been grouped into broader functional categories, which are labelled. Only significantly ($P < 0.05$) over-represented GO categories are shown. A superscript letter R following a GO term description indicates the GO category is over-represented only in the resistant reaction. A superscript letter S following the GO terms description indicates the GO category is over-represented only in the susceptible reaction. RS indicates the term is over-represented in both resistant and susceptible reactions.

process were over-represented only in the susceptible reaction (Fig. 1).

To elucidate how gene expression and function are related to plant defence, we focussed on the GO biological process terms related to defence, stress, hormone signalling, transcription, translation and photosynthesis (Fig. 1). Five GO terms related to defence or stress response were significantly over-represented only in the resistant reaction (Fig. 1). Looking specifically at the GO term defence response, we identified 102 differentially expressed genes (DEGs) including lipoxygenases, disease resistance proteins and others. Only responses to wounding and karrikin were over-represented in both the susceptible and resistant reactions. For these GO terms, the number of DEGs was similar for both reactions. When we examined hormone signalling, only response to salicylic acid was uniquely over-represented in the resistant reaction (Fig. 1). Although the GO terms photosynthesis and photosynthesis light reaction were over-represented in both resistant and susceptible reactions, photosynthetic electron transport and chloroplast organisation were uniquely over-represented in the susceptible reaction.

Two GO terms related to transcription and translation (regulation of transcription and regulation of defence response) were over-represented only in the resistant reaction. For these terms, the number of DEGs identified in the resistant reaction was higher than in the susceptible reaction. Translation and ribosome biogenesis GO terms were over-represented only in the susceptible reaction. Although almost all GO terms had less than a 2-fold difference in gene number between resistant and susceptible reactions, the GO term translation had seven times more genes identified in the susceptible reaction.

As we examined the data in Fig. 1, it became clear that genes common to both reactions might have different expression patterns than those genes unique to each reaction. To better understand these differences, the dataset was divided into three categories, DEGs unique to resistance, common to resistant and susceptible interactions and unique to susceptibility (2337, 1953 and 2144 genes respectively). We then examined the expression of genes within each of the significantly over-represented GO terms (Fig. 2).

Genes common to both the resistant and susceptible reactions exhibited very similar expression patterns (Fig. 2). Genes that were induced at 12 hai were not differentially expressed at 24 hai but had an increase in gene expression at 72 hai. In contrast, genes that were repressed at 12 hai, were not significantly differentially expressed at 24, 72 and 144 hai, but were repressed again at 216 hai infection. This suggests there is a difference in the biphasic response depending on the initial expression of the genes. For genes induced at 12 hai, the lack of

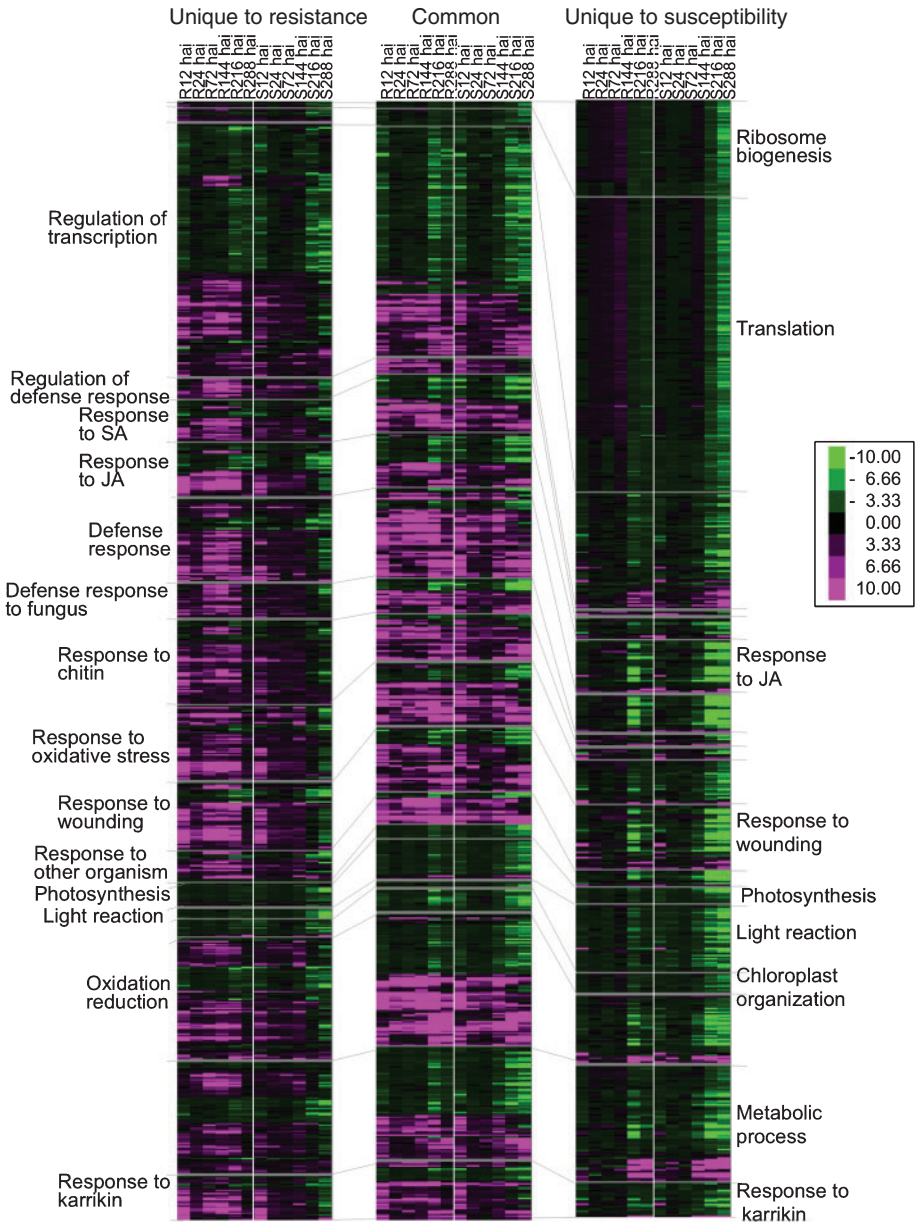

Fig. 2. Analysis of gene expression across the *Rpp4* time course experiment and by gene ontology biological function category. Differences in biphasic gene expression depend on the initial direction of expression of differentially expressed genes; induced genes have a shorter lag than repressed genes. To aid in interpretation, the data were divided into three categories: probe sets differentially expressed only during the resistant reaction, probe sets differentially expressed in both the resistant and susceptible reaction (common) and probe sets differentially expressed only in the susceptible reaction. The expression of a particular probe set is shown for six time points in the resistant reaction and six time points in the susceptible reaction. To put genes in a biological context, only differentially expressed probe sets mapping to an over-represented GO biological process category (Fig. 1) are shown. The statistical analysis of GO data was corrected for multiple probe sets per gene (Fig. 1) and the heatmaps show the expression of all probe sets assigned to a GO category. Data is shown for all 19 over-represented GO categories shown in Fig. 1 and the order of the GO categories is the same across all three panels. Hierarchical clustering was used to group probe sets with similar expression patterns within a GO category. GO categories significantly over-represented in the resistant reaction are labelled on the left panel, whereas GO categories over-represented in the susceptible reaction are labelled on the right panel. Grey lines between panels link GO categories across the uniquely resistant, common and uniquely susceptible datasets. Black boxes represent no change in gene expression compared with the uninfected mock-inoculated treatment, magenta indicates upregulation by *Phakopsora pachyrhizi* inoculation, and green indicates downregulation. More intense colours represent greater fold change (log2 transformed), as shown on the scale.

differential expression spanned a single time point. However, this lag lasted three time points in genes that were initially repressed. We noticed that most genes associated with defence responses were induced while genes associated with translation and photosynthesis were repressed. For the GO terms regulation of transcription and metabolism, equal numbers of genes were repressed and induced in both resistant and susceptible reactions.

Next we examined DEGs unique to the resistant reaction. We included the expression of the same genes in the susceptible reaction (though not statistically significant) to see if similar trends occurred in both reactions. The genes unique to the resistant reaction had the same expression patterns observed above for the common genes, however, in the susceptible reaction, the differential expression of all genes as increasingly suppressed at all time points after 12 hai (Fig. 2).

Finally, DEGs unique to the susceptible reaction were examined. Again, the expression of the same genes in the resistant reaction was included (though not statistically significant) to see whether similar trends were observed in both reactions. Results showed that the expression of almost all genes unique to the susceptible reaction and across all GO terms was repressed (Fig. 2). At 12 hai, we observed mild repression, with little differential expression observed at 24, 72 and 144 hai. Maximum repression occurred at 216 and 288 hai. In contrast, maximum repression occurred earlier in the resistant interaction. Also noteworthy, there was a slight induction at 24, 48 and 72 hai of genes involved in ribosome biogenesis and translation observed only in the resistant reaction; however, this was not statistically significant.

Of all the over-represented GO biological process terms identified, regulation of transcription was the largest GO term identified, containing 300 DEGs from the resistant reaction and 241 DEGs from the susceptible reaction. The difference in expression patterns between the uniquely resistant and susceptible DEGs suggested that transcription factors play a major role in regulating responses to *P. pachyrhizi.* To test if particular transcription factor classes were significantly ($P < 0.05$) differentially expressed, we used the SoyDB soybean transcription factor database (Wang *et al*. 2010) to identify transcription factors and classes present on the array. Fisher's exact test (Fisher 1966) with a Bonferroni correction (Bonferroni 1935) was used to identify over-represented transcription factor classes. In the resistant reaction, 47 different transcription factor classes were represented within the DEGs. Of these, only the WRKY and AUX-IAA-ARF classes were significantly over-represented (Fig. 3). In the susceptible reaction, 43 different transcription factor classes were differentially expressed, but only the NAC, C2C2 (Zn) CO-like and MYB/HD transcription factor classes were significantly over-represented. We note that the DDT, EIL, HMG, PLATZ, R3H and SRS transcription factor classes were found only in the resistant reaction, whereas the HTH-FIS, SNF2 and TUB transcription factor classes were found only in the susceptible reaction.

In addition to identifying transcription factors in the *Rpp4* time course dataset, the expression of the transcription factors themselves was also examined (Fig. 3). The transcription factors were divided into three classes: unique to the resistant reaction, common to both reactions and unique to the susceptible reaction. When compared with Fig. 2, which shows expression of genes from over-represented biological process terms, some obvious differences in expression patterns were observed. Although many downstream genes are induced by *P. pachyrhizi* inoculation (Fig. 2), the majority of transcription factors are repressed (Fig. 3). However, representatives of transcription factor families previously associated with defence and abiotic stress (WRKY, MYB and NAC) were induced. When subdividing the transcription factor data into three classes, we observed the same biphasic responses observed for downstream genes, with different biphasic responses for induced and repressed transcription factors. One notable difference was that both MYB and NAC transcription factors show a biphasic response across all three comparisons, while the WRKY transcription factors unique to resistance had increased expression over time.

Comparison of resistance responses governed by Rpp2*,* Rpp3 *and* Rpp4

Our group has focussed primarily on the reactions mediated by specific *Rpp* genes to *P. pachyrhizi* (van de Mortel *et al*. 2007; Schneider *et al*. 2011). However, little is known on the overlap of molecular responses to *P. pachyrhizi* conditioned by different *R*-genes and different *P. pachyrhizi* isolates. van de Mortel *et al*. (2007) measured gene expression over a 7-day time course in mock-inoculated and inoculated leaves of soybean accession PI230970 (*Rpp2*) and the susceptible line Embrapa-48 infected with a Brazilian field isolate of *P. pachyrhizi.* Schneider *et al*. (2011) measured gene expression in leaves over 12 days in a single soybean accession (PI46312 (*Rpp3*)) inoculated with either incompatible (HW94–1) or compatible (TW80–2) *P. pachyrhizi* isolates. Here, we measured gene expression over 12 days using soybean accession PI450925B (*Rpp4*) and cultivar Wm82, inoculated with *P. pachyrhizi* isolate (HW94–1). These experiments identified 894, 8447, and 5805 differentially expressed probe sets associated with the resistance response governed by *Rpp2*, *Rpp3* and *Rpp4* respectively. Similarly, 1516, 1827 and 5304 differentially expressed probe sets were associated with the susceptible responses, respectively. Comparison of these datasets identified 214 probe sets common to all compatible and incompatible *P. pachyrhizi*-soybean reactions (Table S3) and 54 probe sets specific to *P. pachyrhizi* resistance reactions governed by *Rpp2*, *Rpp3* and *Rpp4* (Table S4). The first comparison could aid in the identification of genes involved in basal defence and the second could identify genes specific to *R*-gene mediated resistance responses. In addition, we were investigated transcription factors and receptors that regulate the expression of downstream defence genes. Of the 214 probe sets common to all resistant and susceptible *P. pachyrhizi* interactions, 22 probe sets (corresponding to 14 unique genes) encoded potential transcription factors or receptors. Of these, six genes (11 probe sets) had best hits to *Arabidopsis* proteins with known roles in basal defence or abiotic stress (Fig. 4*a*) including homologues of AtERF1, AtGATA3, AtWRKY31, AtWRKY33, AtWRKY40 and

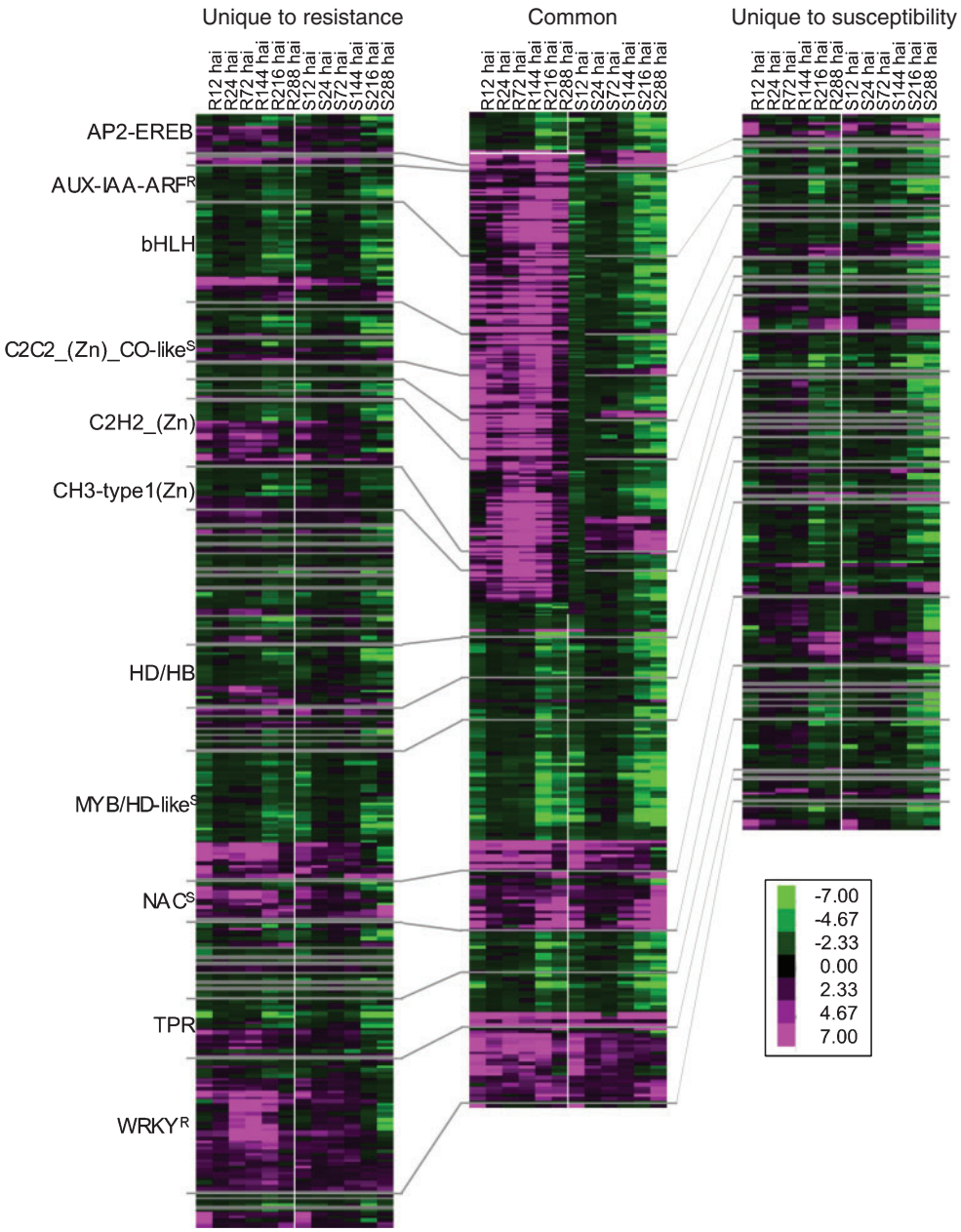

Fig. 3. Significantly differentially expressed soybean transcription factors also demonstrate biphasic responses to *Phakopsora pachyrhizi* infection. The SoyDB website (Wang *et al*. 2010) was used to identify all transcription factors on the Soybean GeneChip Genome Array. Expression data is shown for all significantly differentially expressed transcription factors; however, only large families are labelled. Expression of all probe sets is shown for transcription factors represented by multiple probe sets. To allow comparisons with Fig. 2, the data were divided into three groups: differentially expressed transcription factors unique to the resistant reaction, common to resistant and susceptible reactions and unique to the susceptible reaction. Hierarchical clustering was used to group differentially expressed probe sets with similar expression patterns within a transcription factor family. Colouring and data labels are the same as in Fig. 1. A superscript letter R following a transcription factor family indicates it is significantly over-represented only in the resistant reaction. A superscript letter S following indicates the transcription factor family is significantly over-represented only in the susceptible reaction.

AtWRKY75. Of the 54 probe sets specific to resistance responses governed by *Rpp2*, *Rpp3* and *Rpp4*, 16 probe sets (corresponding to 14 unique genes) encoded potential signalling proteins (Fig. 4*b*). Nine genes (10 probe sets) had best hits to *Arabidopsis* proteins with known roles in defence or abiotic stress including AtL6, AtMPK3, AtSZF2, AtWRKY11, AtWRKY33, AtWRKY40, AtWRKY51, AtWRKY53 and AtZAT10.

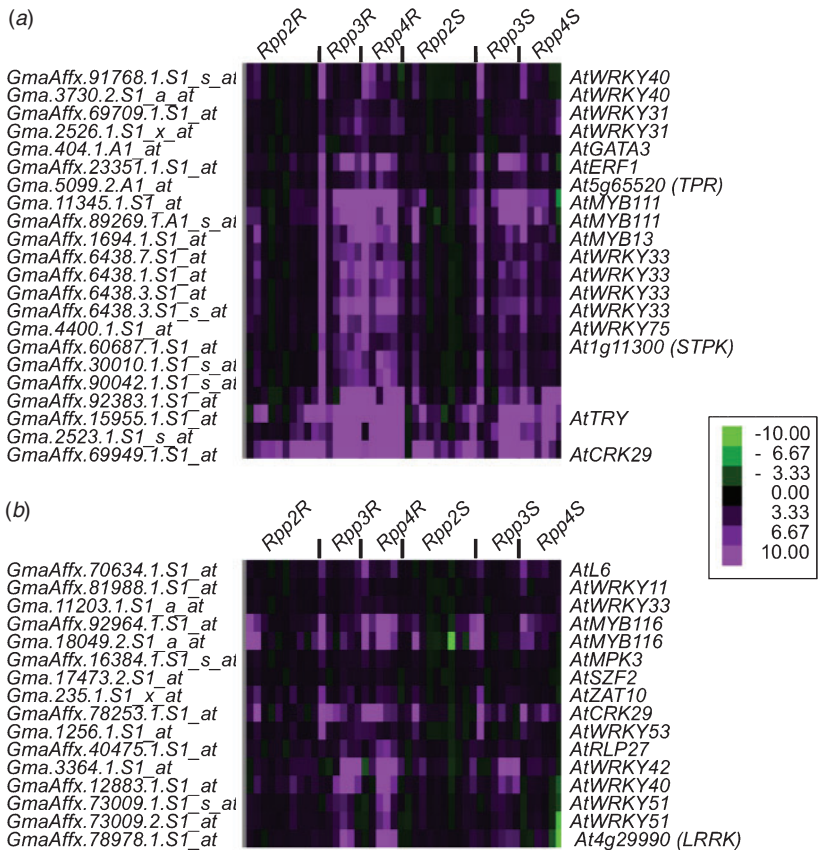

Fig. 4. Transcription factors common to resistance and susceptibility are expressed differently to those unique to resistance. (*a*) Transcription factors significantly differentially expressed and common to resistant and susceptible reactions across multiple microarray experiments were identified. Probe sets were significantly differentially expressed in the *Rpp4* time course described here (PI459025B (*Rpp4*R) and Wm82 (*Rpp4*S)), the *Rpp3* time course described by Schneider *et al*. 2011 (PI462312 (*Rpp3*R and *Rpp3*S)) and the time course described by van de Mortel *et al*. (2007) (PI230970 (*Rpp2*R) and Embrapa-48 (*Rpp2*S). The time points for the *Rpp3* and *Rpp4* experiments include 12, 24, 72, 144, 216 and 288 hours after inoculation (hai). The time points for the van de Mortel *et al*. (2007) (*Rpp2*R and *Rpp2*S) include 6, 12, 18, 24, 36, 48, 72, 96, 120 and 168 hai. (*b*) Transcription factors significantly differentially expressed only in resistance reactions across multiple microarray experiments were identified (*Rpp2*R, *Rpp3*R and *Rpp4*R). Data for the susceptible reactions are shown for reference. Black boxes represent no change in gene expression compared with the uninfected mock-inoculated treatment, magenta indicates upregulation by *Phakopsora pachyrhizi* inoculation, and green indicates downregulation. More intense colours represent greater fold change (log2 transformed), as shown on the scale. Hierarchical clustering was used to group probe sets with similar expression patterns. The probe identifiers are shown on the left of the panel and the best *Arabidopsis* matches are shown on the right.

Comparison of the signalling genes conserved across all resistant and susceptible reactions to those signalling genes specific to resistant responses, revealed obvious differences in expression pattern. Although the majority of signalling genes were induced by *P. pachyrhizi* infection, genes differentially expressed in both resistant and susceptible reactions tend to have broader and greater expression across time points (Fig. 4*a*, *b*). We noted that soybean homologous of *AtWKRY33* (Glyma11 g29720 and Glyma02 g38970), *AtWRKY40* (Glyma08 g23380 and Glyma17 g33890) and *AtCRK29* (Glyma08 g46650 and Glyma20 g27480) were found in both datasets. In all cases, each gene in the homologous pair had a best match against the same *Arabidopsis* protein, yet homologous pairs had very different expression patterns. None of the pairs identified included homeologous genes derived from soybean's duplicated genome (data not shown).

Gene Expression in Rpp4*-silenced and empty vector treated PI459025B plants, following inoculation with* P. pachyrhizi

The *Rpp4* time course experiment led to the identification of soybean genes that are differentially expressed in response to

P. pachyrhizi infection. However, not all of these genes are downstream of the *Rpp4* gene. We used RNA isolated from the same plants described by Meyer *et al*. (2009) to identify genes downstream of *Rpp4* in the resistance response. RNA from PI459025B plants pre-treated with either the BPMV-LRR construct, which caused loss of *Rpp4* resistance, or the BPMV empty vector control and subsequently inoculated with *P. pachyrhizi* isolate LA04–1, which induces a resistance response in PI459025B. By comparing RNA from *Rpp4*-silenced (BPMV-LRR) and non-silenced (BPMV empty vector) PI459025B plants, we expected to identify genes downstream of *Rpp4* in the signalling pathway controlling resistance to *P. pachyrhizi*. RNA samples were hybridised to the Gene Chip Soybean Genome Array (Affymetrix) and analysed for differential gene expression. Results were deposited in the Plant Expression Database (accession number GM51; Wise *et al*. 2007).

Using this approach, a total of 383 significantly ($P < 0.05$) differentially expressed probe sets, representing 264 unique genes, were identified with a 2-fold difference in expression between *Rpp4*-silenced and non-silenced plants (Table S5). Although 19 DEGs were induced in *Rpp4*-silenced plants relative to non-silenced plants, most (245 DEGs) were repressed. The same annotation pipeline described earlier was used to characterise the identified DEGs. Table 1 details the expression and annotation of the 10 most induced and repressed genes identified. Several well characterised defence genes were repressed by silencing of *Rpp4* including pathogenesis-related protein PR1a and genes related to flavonoid (chalcone synthase and flavonol synthase), phenylpropanoid and lignin (caffeoyl-CoA o-methyltransferase2), ethylene (1-aminocyclopropane-1-carboxylate oxidase) and jasmonate synthesis (12-oxophytodienoic acid 10). In addition, a homologue of the photoassimilate-responsive protein PAR was also repressed by *Rpp4* silencing. Genes induced by the silencing of *Rpp4* included pectin acetylesterase, a homologue of pepsin A and an auxin-responsive protein (Table 1).

To understand the pathways affected by *Rpp4* silencing and to place the DEGs in a biological context, we examined the GO biological process terms (Ashburner *et al*. 2000) to identify any

Table 1. Annotation and expression of the top 10 repressed and induced genes identified by comparison of LRR-silenced and vector treated PI459025B plants following infection with *Phakopsora pachyrhizi*

Affymetrix probe ID[A]	Fold change (LRR/VO)[B]	Glyma1 gene call[C]	Best uniRef100 Hit[D]	UniRef100 e-value
GmaAffx.93635.1.S1_s_at	–11.26	Glyma13 g32540	Q9XFB4 Pathogenesis related protein PR1a, *G. max*	5.00E-97
GmaAffx.77637.1.S1_at	–9.83	Glyma19 g27930	Q2ENC4 Chalcone synthase, *Populus alba*	0.00E+00
Gma.10150.1.A1_at	–8.99	Glyma15 g16490	B9RT28 Flavonol synthase/flavanone 3-hydroxylase, *Ricinus communis*	1.00E-163
GmaAffx.92564.1.S1_at	–8.59	Glyma12 g05780	A8TVQ5 Beta-glucosidase G2, *Medicago truncatula*	0.00E+00
GmaAffx.92558.1.S1_s_at	–7.51	Glyma16 g01990	B9RMV2 1-aminocyclopropane-1-carboxylate oxidase, *Ricinus communis*	1.00E-148
Gma.14338.1.A1_at	–7.28	Glyma05 g30290	D3W0X7 Globulin, *Vitis berlandieri* × *Vitis riparia*	1.00E-119
Gma.3604.4.S1_s_at	–6.75	Glyma11 g05800	Q9SWB8 Caffeoyl-CoA O-methyltransferase 2, *Eucalyptus globulus*	1.00E-102
GmaAffx.57966.1.S1_at	–6.55	Glyma09 g25420	Q8 LDW1 Photoassimilate-responsive protein PAR-like protei, *Arabidopsis thaliana*	5.00E-50
GmaAffx.18868.1.S1_s_at	–6.41	Glyma15 g35410	C5I848 12-oxophytodienoic acid 10,10-reductase, *Astragalus sinicus*	1.00E-151
Gma.2586.1.S1_at	–6.20	Glyma08 g09070		
GmaAffx.61395.1.A1_at	2.36	Glyma13 g31600		
Gma.2961.1.S1_at	2.46	Glyma08 g19000	A6XNC5 (Iso)flavonoid glycosyltransferase, *Medicago truncatula*	1.00E-157
Gma.6498.1.A1_at	2.54	Glyma11 g03500	B9SSF8 Pepsin A, putative, *Ricinus communis*	1.00E-148
GmaAffx.53274.1.S1_at	2.58	Glyma03 g33110	B7FN40 Putative uncharacterised protein, *Medicago truncatula*	3.00E-19
Gma.7454.1.S1_a_at	2.65	Glyma13 g43800	B9RUW0 Auxin-responsive protein IAA1, *Ricinus communis*	4.00E-40
Gma.1007.2.S1_at	2.78	Glyma11 g33790	Q9AR93 Putative calmodulin-related protein, *Medicago truncatula*	7.00E-58
GmaAffx.4935.2.S1_at	2.79	Glyma10 g28230	Q2HRU2 Pectinacetylesterase, *Medicago truncatula*	1.00E-172
GmaAffx.48606.1.S1_at	3.25	Glyma13 g33780		
GmaAffx.4935.1.S1_at	4.09	Glyma20 g22210	Q2HRU2 Pectinacetylesterase, *Medicago truncatula*	0.00E+00
Gma.4755.1.S1_at	4.39	Glyma17 g17210		

[A]Affymetrix soybean GeneChip array ID.

[B]Fold change in gene expression between leaves of LRR silenced and vector only plants infected with *Phakopsora pachyrhizi*.

[C]Matching soybean gene. Affymetrix soybean target sequences were compared with predicted Glyma1 cDNAs (Schmutz *et al*. 2010) using BlastN (Altschul *et al*. 1997) and requiring 95% nucleotide identity and $E < 10^{-30}$.

[D]Best uniref sequence match. If a matching Glyma1 protein sequence was identified, this was used for BLASTP against Uniref100 (ver. 12/2010, Apweiler *et al*. 2004), otherwise the Affymetrix soybean consensus sequence was compared using BLASTX. Only data for the best high scoring pair is reported ($E < 10^{-6}$).

Table 2. Gene ontology biological process terms (Ashburner *et al*. 2000) significantly ($P < 0.05$) over-represented among differentially expressed genes identified by comparison of LRR-silenced and vector treated PI459025B plants following infection with *Phakopsora pachyrhizi*

Before statistical analysis, differentially expressed probe sets were mapped to unique soybean genes. In the statistical analysis, only gene counts were used to avoid inflation caused by multiple probe sets per gene

Gene ontology biological process term	Gene array count	Differentially expressed gene count	Corrected *P*-value	Gene ontology biological process description
GO:0009813	53	13	3.88E-08	Flavonoid biosynthetic process
GO:0010422	6	5	7.76E-06	Regulation of brassinosteroid biosynthetic process
GO:0016131	8	5	6.97E-05	Brassinosteroid metabolic process
GO:0009809	63	10	3.72E-04	Lignin biosynthetic process
GO:0051555	36	7	3.30E-03	Flavonol biosynthetic process
GO:0009699	25	6	3.72E-03	Phenylpropanoid biosynthetic process
GO:0006979	265	18	1.16E-02	Response to oxidative stress
GO:0010224	66	8	2.88E-02	Response to UV-B

that were significantly over-represented in the VIGS dataset when compared with all genes on the soybean array as described earlier. We identified eight significantly over-represented biological process terms associated with general defence, brassinosteroid signalling and phenylpropanoid production (Table 2).

Identification of significantly over-represented transcription factor binding sites in the promoters of genes differentially expressed in response to Rpp4*-silencing*

Analysis of the GO biological process terms from the resistant and susceptible time course experiment revealed that regulation of transcription was the most frequently identified GO term. In order to understand how differentially expressed genes in the *Rpp4*-signalling pathway were regulated, we wanted to identify differentially expressed transcription factors in the *Rpp4* VIGS dataset. Transcription factors were identified using the SoyDB transcription factor database (Wang *et al*. 2010). Annotation of the differentially expressed genes identified 12 repressed transcription factors representing the AP2-EREBP, BZIP, C2H2 (Zn), CCAAT, Homeodomain/HOMEOBOX, HSF, MYB/HD-like and TPR transcription factor families (Table 3). Of the three MYB/HD-like transcription factors identified, one was a distant homologue of *GmMYB84*, required for *Rpp2*-mediated defence against *P. pachyrhizi* (Pandey *et al*. 2011). Seven of the transcription factors identified have not been previously associated with stress or defence responses. In addition, two transcription factors induced in *Rpp4*-silenced plants including an SRS and an AUX-IAA-ARF transcription factor were identified.

Our microarray analyses of the VIGS-treated plants was limited to a single time point late in the infection process (14 dai) making it difficult to identify transcription factors that operate early in the *Rpp4*-signalling pathway. Nevertheless, we used Clover (Frith *et al*. 2004) to identify transcription factor binding sites (TFBS) significantly ($P < 0.05$) over-represented in the promoters of genes corresponding to a particular GO category (Table 2) compared with all promoters in the soybean genome. The analysis was limited to plant TFBS in the TRANSFAC transcription factor database (Matys *et al*. 2006) and 1000 bp of promoter sequence. Our initial analyses of differentially expressed transcription factors identified three MYB transcription factors (Table 3). Clover analyses identified five different MYB TFBS associated with over-represented biological process terms including TFBS for the MYBAS1, AtMYB-84, AtMYB-15, C1 and P transcription factors (Tables 4, S6). Two differentially expressed BZIP transcription factors were also identified that were homologues of AtBZIP9, which has not been previously associated with stress or defence responses. However, we found several BZIP TFBS over-represented within the GO biological process terms including OSBZ8, TGA1, ABI4, ABZ1, ABF and TRAB1, which have known roles in stress and defence responses (Table 4). Previously, GmMYB84 and GmWRKY36 were shown to be required for *Rpp2*-mediated resistance (Pandey *et al*. 2011). Although MYB TFBS were associated with all of the GO biological process terms, the WRKY TFBS were specific to phenylpropanoid biosynthesis and response to UV-B (Tables 4, S6).

If differentially expressed genes were controlled by *Rpp4*, it was expected that genes normally induced in an *Rpp4*-genetic background would be repressed by *Rpp4*-silencing and similarly, genes normally repressed would be induced by silencing. However, the time points of the *Rpp4* time course experiment were different than the single time point used for the VIGS experiment. Therefore, we compared the expression of probe sets common to the incompatible time course reaction and the VIGS experiment. Of the 384 probe sets identified in the VIGS experiment, 235 (162 unique genes) were also differentially expressed in the incompatible reaction of the time course experiment (Fig. 5). As predicted, genes normally induced during *Rpp4*-mediated resistance were repressed by *Rpp4* silencing. Six probe sets, corresponding to six unique genes, were repressed during *Rpp4*-mediated resistance but induced by *Rpp4* silencing.

We also compared the differentially expressed genes identified in the VIGS dataset to all differentially expressed genes in the *Rpp2* (van de Mortel *et al*. 2007), *Rpp3* (Schneider *et al*. 2011) and *Rpp4* datasets. We identified 101 probe sets, corresponding to 72 genes, unique to the *Rpp4* VIGS dataset. GO biological process analysis within these 72 unique genes demonstrated that GO categories nucleosome assembly

Table 3. Significantly differentially expressed transcription factors in the *Rpp4* signalling pathway

Affymetrix probe ID[A]	Fold change[B]	Glyma1 gene call[C]	TF family[D]	Best *Arabidopsis* hit[E]	E-value	Other *Arabidopsis* name	Defence or stress reference
GmaAffx.61439.1.A1_at	–3.43	Glyma19 g34220	C2H2 (Zn)	AT1G03840	3.00E-89	AtIDD2	
GmaAffx.21313.3.S1_at	–3.29	Glyma04 g08290	BZIP	AT5G24800	2.00E-28	ATBZIP9	
GmaAffx.89269.1.A1_s_at	–3.02	Glyma16 g02570	MYB/HD-like	AT1G22640	5.00E-52	ATMYB3	Cheong *et al.* (2002)
GmaAffx.21313.1.S1_s_at	–2.79	Glyma06 g08390	BZIP	AT5G24800	1.00E-30	ATBZIP9	
GmaAffx.92200.1.S1_at	–2.79	Glyma06 g08390	BZIP	AT5G24800	1.00E-30	ATBZIP9	
GmaAffx.132.1.S1_at	–2.69	Glyma13 g30720	AP2-EREBP	AT5G47220	3.83E-03	ATERF2	Ohme-Takagi and Shinshi (2000)
Gma.17063.1.S1_at	–2.66	Glyma04 g08290	BZIP	AT5G24800	2.00E-28	ATBZIP9	
GmaAffx.7138.1.S1_at	–2.62	Glyma17 g16980	MYB/HD-like	AT3G49690	1.00E-61	ATMYB84	Pandey *et al.* (2011)
Gma.16196.1.S1_at	–2.59	Glyma12 g36540	CCAAT	AT5G12840	4.00E-25	ATHAP2A	
Gma.11345.1.S1_at	–2.57	Glyma16 g02570	MYB/HD-like	AT1G22640	5.00E-52	ATMYB3	Cheong *et al.* (2002)
GmaAffx.28690.1.S1_at	–2.35	Glyma06 g01940	Homeodomain/ HOMEOBOX	AT4G35550	1.00E-53	WOX13	
GmaAffx.92807.1.S1_at	–2.26	Glyma07 g05960	MYB/HD-like	AT5G49330	7.00E-54	ATMYB11	Stracke *et al.* (2010)
GmaAffx.1890.1.S1_s_at	–2.14	Glyma05 g20460	HSF	AT4G36990	3.00E-14	ATHSF4	
GmaAffx.39668.1.S1_at	–2.13	Glyma05 g20460	HSF	AT4G36990	3.00E-14	ATHSF4	
GmaAffx.54770.1.S1_at	–2.09	Glyma10 g43630	C2H2 (Zn)	AT2G37430	1.00E-23	ZAT11	
GmaAffx.85020.1.S1_at	–2.09	Glyma13 g34610	TPR	AT1G53300	0.00E+00	TTL1	Lakhssassi *et al.* (2012)
Gma.5599.1.A1_at	2.00	Glyma14 g03900	SRS	AT5G12330	5.00E-56	LRP1	
Gma.7454.1.S1_a_at	2.65	Glyma13 g43800	AUX-IAA-ARF	AT1G04250	4.00E-35	IAA17	

[A]Affymetrix soybean GeneChip array ID.

[B]Fold change in gene expression between leaves of LRR silenced and vector only plants infected with *Phakopsora pachyrhizi*.

[C]Matching soybean gene. Affymetrix soybean target sequences were compared with predicted Glyma1 cDNAs (Schmutz *et al.* 2010) using BlastN (Altschul *et al.* 1997) and requiring 95% nucleotide identity and $E < 10^{-30}$.

[D]Transcription factor families were identified using the SoyDB transcription factor database (Wang *et al.* 2010).

[E]The best *Arabidopsis thaliana* match was identified by BLASTP against all predicted *A. thaliana* proteins. If a matching Glyma1 protein sequence was identified, this was used for BLASTP against Uniref100 (ver. 12/2010, Apweiler *et al.* 2004). Otherwise, the Affymetrix Soybean consensus sequence was compared using BLASTX. Only data for the best high scoring pair is reported ($E < 10^{-6}$).

(GO:0006334, *P* <3.4E-07) and phenylpropanoid biosynthesis (GO:0009699, *P* <3.7E-03) were significantly over-represented. Additionally, this dataset included five genes (Glyma04 g08290, Glyma06 g08390, Glyma13 g34610, Glyma13 g43800 and Glyma19 g34220) with best matches to *Arabidopsis* transcription factors AtBZIP9, AtTTL1, AtIAA17 and AtIDD2. All of these genes were repressed by *Rpp4*-silencing except the homologue of AtIAA17 (Glyma13 g43800).

Discussion

Time-course analysis identified two patterns in biphasic soybean responses to P. pachyrhizi *infection*

We have used microarray expression profiling, VIGS and the soybean genome sequence (Schmutz *et al.* 2010) to begin developing a genetic framework for *Rpp4*-mediated resistance against *P. pachyrhizi*. Previously, we reported on gene expression changes occurring in plants carrying the *Rpp2* and *Rpp3* resistance genes (van de Mortel *et al.* 2007; Schneider *et al.* 2011). The resistance phenotypes of *Rpp2* and *Rpp3* are similar to *Rpp4*, yet all three genes recognise distinct isolates of *P. pachyrhizi.* To allow the eventual comparison of gene expression governed by the *Rpp3* and *Rpp4* resistance genes, experiments were conducted at the same facility, using the same inoculation protocols and time courses. Further, the same statistical and annotation methods were used to identify differentially expressed genes. In the *Rpp3* experiment, we compared gene expression in a single soybean genotype infected with two different isolates of *P. pachyrhizi*, allowing us to control for the genetic background typically found when comparing two different genotypes. Using this approach, we found that genes differentially expressed in the susceptible response were a subset of those expressed in the resistant response. In the *Rpp4* time-course described here, we used two different soybean lines infected with the same *P. pachyrhizi* isolate. These different approaches affect how we examine and interpret these datasets.

In the case of the *Rpp3* (Schneider *et al.* 2011) and *Rpp4* microarray time course experiments, we chose to divide the differentially expressed genes into three categories: unique to resistance, unique to susceptibility and common to resistance and susceptibility. When we looked at the expression patterns of genes of common to resistance and susceptibility, the majority of genes in the *Rpp3* experiment had a biphasic expression pattern and were induced by *P. pachyrhizi* infection. In the resistant reaction, differential gene expression at 12 hai was muted until 72 hai. In the susceptible reaction, differential gene expression was muted until 144 hai. In contrast, the *Rpp4* time course experiment described here showed that approximately half of the genes common to the resistant and susceptible reactions were repressed across all time points, whereas the remaining genes were induced. Further, the

Table 4. Identification of significantly over-represented transcription factor binding sites (TFBS) in the promoters of genes associated with significantly over-represented GO biological process terms

Table S4 provides a complete list of all over-represented plant TFBS and the corresponding motif identifiers

GO biological process term[A]	GO description	Number of over-represented TFBS[B]	Defence or stress associated TFBS[C]	TF name	TF class	Clover *P*-value ($P < 0.05$)	Reference
GO:0006979	Response to oxidative stress	13	2	OSBZ8	BZIP	0.004	Mukherjee *et al*. (2006)
				MYBAS1	MYB	0.008	Yang *et al*. (2001)
GO:0009699	Phenylpropanoid biosynthetic process	41	16	AtMYB-84	MYB	0.000	Pandey *et al*. (2011)
				ZAP1	WRKY	0.001	Assunção *et al*. (2010)
				AtMYB-15	MYB	0.003	Dubos *et al*. (2010)
				TGA1b	BZIP	0.007	Lindermayr *et al*. (2010)
				ABI4	BZIP	0.004	Giraud *et al*. (2009)
				TGA1a	BZIP	0.004	Lindermayr *et al*. (2010)
				LIM1	LIM	0.004	Kawaoka and Ebinuma (2001)
				ABZ1	BZIP	0.007	Sell and Hehl (2004)
				NAC69–1	NAC	0.013	Xue *et al*. (2011)
				WRKY	WRKY	0.024	Rushton *et al*. (2010)
				P	MYB	0.030	Grotewold *et al*. (1994)
				ABF	BZIP	0.037	Kang *et al*. (2002)
				TRAB1	BZIP	0.040	Agarwal and Jha (2010)
				ABF1	BZIP	0.045	Kang *et al*. (2002)
				C1	MYB	0.045	Dubos *et al*. (2010)
GO:0009809	Lignin biosynthetic process	10	5	LIM1	LIM	0.005	Kawaoka and Ebinuma (2001)
				C1	MYB	0.009	Dubos *et al*. (2010)
				Alfin1	PHD	0.022	Wei *et al*. (2009)
				AtMYB-84	MYB	0.023	Pandey *et al*. (2011)
				OSBZ8	BZIP	0.049	Mukherjee *et al*. (2006)
GO:0009813	Flavonoid biosynthetic process	17	6	AtMYB-84	MYB	0.001	Pandey *et al*. (2011)
				P	MYB	0.002	Grotewold *et al*. (1994)
				AtMYB-15	MYB	0.023	Dubos *et al*. (2010)
				C1	MYB	0.040	Dubos *et al*. (2010)
				MYBAS1	MYB	0.044	Yang *et al*. (2001)
				TGA1a	BZIP	0.048	Lindermayr *et al*. (2010)
GO:0010224	Response to UV-B	40	15	AtMYB-84	MYB	0.000	Pandey *et al*. (2011)
				P	MYB	0.000	Grotewold *et al*. (1994)
				OSBZ8	BZIP	0.001	Mukherjee *et al*. (2006)
				ABI4	BZIP	0.005	Giraud *et al*. (2009)
				TRAB1	BZIP	0.005	Agarwal and Jha (2010)
				AtMYB-15	MYB	0.006	Dubos *et al*. (2010)
				ZAP1	WRKY	0.006	Assunção *et al*. (2010)
				C1	MYB	0.007	Dubos *et al*. (2010)
				ABF	BZIP	0.009	Kang *et al*. (2002)
				WRKY	WRKY	0.017	Rushton *et al*. (2010)
				ABF1	BZIP	0.027	Kang *et al*. (2002)
				TGA1a	BZIP	0.027	Lindermayr *et al*. (2010)
				LIM1	LIM	0.037	Kawaoka and Ebinuma (2001)
				TGA1b	BZIP	0.020	Lindermayr *et al*. (2010)
GO:0010422	Regulation of brassinosteroid biosynthetic process	3	2	AtMYB-15	MYB	0.022	Dubos *et al*. (2010)
				AtMYB-84	MYB	0.024	Pandey *et al*. (2011)
GO:0016131	Brassinosteroid metabolic process	3	2	AtMYB-15	MYB	0.022	Dubos *et al*. (2010)
				AtMYB-84	MYB	0.024	Pandey *et al*. (2011)

Table 4. (*continued*)

GO biological process term[A]	GO description	Number of over-represented TFBS[B]	Defence or stress associated TFBS[C]	TF name	TF class	Clover *P*-value ($P < 0.05$)	Reference
GO:0051555	Flavonol biosynthetic process	7	4	C1	MYB	0.006	Dubos *et al.* (2010)
				P	MYB	0.020	Grotewold *et al.* (1994)
				AtMYB-84	MYB	0.036	Pandey *et al.* (2011)
				MYBAS1	MYB	0.043	Yang *et al.* (2001)

[A]Gene ontology terms correspond to biological process terms in Table 3. Clover (Frith *et al.* 2004) was used to identify over-represented TFBS in the promoters of the genes with a GO term relative to all promoters in the soybean genome.
[B]Analysis was limited to plant TFBS in the TRANSFAC transcription factor database (Matys *et al.* 2006).
[C]TFs associated with defence were identified through the TRANSFAC database or literature review.

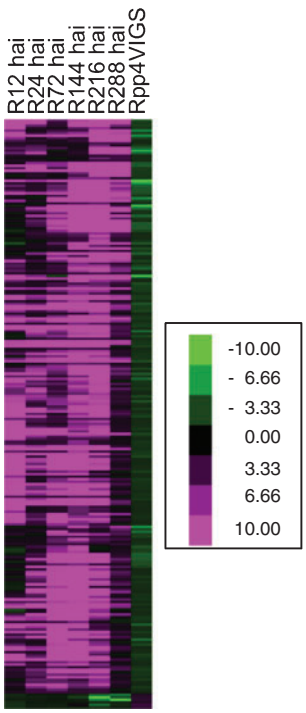

Fig. 5. Identification of genes whose differential expression is dependent upon *Rpp4*. To identify genes dependent on the expression of *Rpp4*, we compared probe sets common to the *Rpp4* resistant reaction time course experiment and the *Rpp4* VIGS experiment. Expression data is shown for all differentially expressed probe sets common to both datasets. Hierarchical clustering was used to group differentially expressed probe sets with similar expression patterns. The first six lanes correspond to time points in the incompatible time-course experiment. The last lane corresponds to the VIGS experiment.

biphasic response differed between repressed and induced genes. Differential gene expression was muted from 24 to 144 hai in repressed genes, but muted only at 24 hai in induced genes.

When we compared differential gene expression in the genes unique to resistant reactions, we observed additional differences between the *Rpp3* and *Rpp4* time course experiments. In the *Rpp3* experiment, the majority of genes unique to resistance were repressed across all time points. Both repressed and induced genes had the same biphasic response described for *Rpp3* common genes above. In contrast, the majority of genes unique to resistance in the *Rpp4* microarray were induced, with the same two biphasic responses described for *Rpp4* common genes above.

We also examined genes unique to susceptibility. In the *Rpp3* experiment, where only a single soybean genotype was used, only 27 differentially expressed probe sets unique to susceptibility were identified, whereas in this experiment in which two soybean genotypes were used, we identified 2686 significantly differentially expressed probes sets (2144 unique genes). When we examined the expression of these genes, almost all of these genes were repressed and instead of a biphasic response, we saw increased repression across time.

Previously, Meyer *et al.* (2009) used qRT–PCR to quantify the expression of *Rpp4* candidate genes at the 12, 24, 72 and 216 hai time points in the same PI459025B and Wm82 RNA samples described here. No significant difference in *Rpp4* candidate gene expression was observed between *P. pachyrhizi*-inoculated and mock-inoculated samples, although a slight biphasic response was observed. Taqman RT–PCR of the samples revealed that *Rpp4*-mediated defence begins sometime between 24 and 72 hai, when fungal RNA accumulation begins to slow. This time line for fungal growth fits nicely with the biphasic pattern seen for induced soybean genes in the *Rpp4* time course experiment. Since *Rpp4* encodes a nucleotide-binding site leucine-rich repeat resistance protein (Meyer *et al.* 2009), it is likely that recognition of *P. pachyrhizi* occurs inside the plant cell as the fungal haustoria develop and secrete proteins into the plant cell to aid in nutrient acquisition (Voegele and Mendgen 2003). This occurs between 24 and 72 hai in PI459025B (*Rpp4*), but it occurs no earlier than 72 hai in PI462312 (*Rpp3*, Schneider *et al.* 2011). Note that both PI459025B and PI462312 were inoculated with the same *P. pachyrhizi* isolate HW94–1. Although there could be slight differences in experiments, it is also possible that HW94–1 is able to overcome basal defence responses in PI459025B more easily, allowing it to develop faster and leading to earlier detection by *Rpp4*.

Biological function of differentially expressed genes identified in response to P. pachyrhizi *infection*

In an effort to categorise the genes identified in the *Rpp4* time course experiment we took advantage of GO biological process

information available from *Arabidopsis*. Unexpectedly, 41% of differentially expressed genes had best hits to *Arabidopsis* genes with no informative GO information. However, when we examined the chip as a whole, 47% of probes lacked informative GO information. This suggests that the functions of many genes in the soybean genome remain to be discovered. Among genes that could be assigned GO information, we could see different patterns of expression dependent on gene function. For example, genes involved in defence and stress responses were induced whereas genes involved in translation, ribosome biogenesis, photosynthesis and the regulation of transcription were largely repressed, similar to *Rpp2* and *Rpp3*-mediated responses (van de Mortel *et al*. 2007; Schneider *et al*. 2011). Although these repressed GO categories may seem unrelated, several studies have reported that defence responses to various pathogens included repression of photosynthesis as plants shift towards non-assimilatory metabolism (Bolton 2009; Major *et al*. 2010). Recently, OsNPR1 from rice has been shown to repress genes involved in photosynthesis and chloroplast transcription and translation (Sugano *et al*. 2010). This suggests a role of OsNPR1, which is a regulator of salicylic acid-mediated defence responses, is to facilitate the redistribution of energy to defence. Note that one the most repressed genes identified in the VIGS microarray dataset was a photo-assimilate responsive gene (Glyma09 g25420) that was repressed 6-fold in response to *Rpp4* silencing.

The role of transcription factors in mediating defence responses to P. pachyrhizi

The subtly different biphasic expression patterns seen for induced and repressed genes in the *Rpp4* time course experiment are likely due to the transcription factors regulating gene expression. We can use the same approach described above to examine the expression of transcription factors common to resistant and susceptible responses and unique to resistance. In both the *Rpp3* and *Rpp4* experiments, we saw roughly equal proportions of transcription factors being induced or repressed. However, in the *Rpp3* experiment, we observed that induced and repressed transcription factors common to resistant and susceptible reactions had the same biphasic response from 24 to 72 hai. When we examined transcription factors unique to resistance, this biphasic response extended until 216 hai. Further, the expression of a majority of these genes was repressed (Schneider *et al*. 2011). In contrast, in the *Rpp4* time course experiment, transcription factor expression was similar to downstream differentially expressed genes. For both the common and resistance-specific transcription factors, we observed different biphasic responses depending on the induction or repression of the transcription factor. Induced transcription factors had a short lag in expression at the 24 hai time point, whereas this lag lasted until the 144 hai time point for repressed transcription factors. Although most genes were induced by pathogen inoculation in both the *Rpp3* and *Rpp4* time course experiments, the majority of transcription factors were repressed, suggesting that a small subset of induced transcription factors, such as MYB and WRKY transcription factors, regulate the expression of defence related genes and the majority of transcription factors repress expression of nonessential genes.

MYB transcription factors can be divided into different classes depending on their structure. All four classes are found in plants, representing the taxon with the highest diversity of MYB proteins (Dubos *et al*. 2010). Recently Pandey *et al*. (2011) tested a VIGS construct targeting *GmMYB84*. Silencing of *GmMYB84* compromised *Rpp2*-mediated resistance. It was hypothesised that GmMYB84 may act as a regulator of genes in the phenylpropanoid pathway and demonstrated that expression of *GmPAL1* was affected in *GmMYB84*-silenced *Rpp2* plants. In our experiments, we identified a *GmPAL1* homologue that was repressed in *Rpp4*-silenced plants, suggesting its expression is dependent on *Rpp4* and it may be required for *Rpp4*-mediated resistance. Further, the GO category phenylpropanoid biosynthesis was significantly ($P < 0.05$) over-represented among differentially expressed genes from *Rpp4*-silenced plants. Clover analyses of the promoters of differentially expressed genes in this GO category identified the AtMYB84 transcription factor-binding site as significantly over-represented when compared with all promoters in the soybean genome. These findings suggest that the *GmMYB84* identified by Pandey *et al*. (2011), or a close homologue, may also be required for *Rpp4*-mediated resistance. It is also worth noting that the promoters of genes in the phenylpropanoid biosynthesis pathway were also significantly over-represented with TFBS for 15 other defence or stress associated transcription factors, suggesting the expression of genes in this pathway are tightly regulated by a complex network of transcription factors.

In our analysis of the *Rpp4* time course experiment, the WRKYs were the only transcription factors that demonstrated different expression patterns when divided into the classes unique to resistance, common to resistance and susceptibility and unique to susceptibility. The WRKYs common to resistance and susceptibility and unique to susceptibility were expressed in a biphasic manner. They were strongly expressed at 12 hai, slightly repressed at 24 hai, but induced again at 72 hai. In contrast, the WRKYs unique to the resistance response were weakly induced at 12 and 24 hai but were strongly induced by 72 hai. This expression pattern suggests this WRKY subset is regulated by *Rpp4*.

WRKY transcription factors form a highly integrated regulatory network that modulate gene expression in defence responses and regulate plant growth and development (Rushton *et al*. 2010). Modification of WRKY expression patterns contributes to the control of various signalling pathways and regulatory networks (Chen *et al*. 2012). Silencing of *GmWRKY36* compromised the expression of *GmFMO* (flavine mono-oxygenase), *GmO-MT* and *GmPR1* (Pandey *et al*. 2011), presumably by regulating the expression of these and other genes during *P. pachyrhizi* infection. Promoters of genes from the VIGS dataset involved in the GO categories phenylpropanoid biosynthesis and response to UVB, were over-represented with a WRKY transcription factor-binding site. The upregulation of phenylpropanoid-related pathways and lignin fortification of cell walls are crucial defence mechanisms in *R*-gene mediated resistance (Hückelhoven 2007). Our findings in the *Rpp4* time course and VIGS experiments suggest that *GmWRKY36* and

other WRKY homologues are also required for *Rpp4*-mediated resistance.

The final transcription factor family significantly over-represented in the resistance reaction governed by *Rpp4* was the AUX-IAA-ARF family. Unlike the WRKY and MYB families, the expression of AUX-IAA-ARF transcription factors was repressed across all comparisons, with similar biphasic responses. The repression of these transcription factors in response to *P. pachyrhizi* is consistent with the idea that growth and development are negatively affected during pathogen defence (Bolton 2009). These data are also consistent with a recent study in which silencing of soybean *MAP kinase 4* (*GmMPK4*) was shown to induce constitutive defence responses and repress expression of *Aux/IAA* and *ARF* transcription factor genes (Liu *et al*. 2011). The different transcription factor expression patterns suggest that resistance to *P. pachyrhizi* depends on complex network of transcription factors regulating defence gene networks.

In addition to identifying transcription factors important for *Rpp4*-mediated signalling, our work also allowed identification of transcription factors differentially expressed across all resistant and susceptible time course experiments (van de Mortel *et al*. 2007; Schneider *et al*. 2011). By comparing across resistant and susceptible reactions in multiple *R*-gene backgrounds, we hoped to identify common transcription factors involved in basal defence responses. Differentially expressed transcription factors included soybean homologues of *AtMYB111*, *AtWRKY31*, *AtWRKY33*, *AtWRKY40* and *AtWRKY75*. AtWRKY33, AtWRKY40 and AtWRKY75 have demonstrated roles in defence and resistance (Fig. 4*a*). AtWRKY33 is essential for defence against the necrotrophic fungus *Botrytis cinerea* (Birkenbihl *et al*. 2012). Global expression profiling that compared susceptible *AtWRKY33* mutants and resistant wild-type plants uncovered large differential transcriptional reprogramming upon *B. cinerea* infection. Kinetic analyses revealed that loss of *WRKY33* function results in inappropriate activation of the salicylic acid (SA)-related host response and elevated SA levels post infection. This leads to the downregulation of jasmonic acid (JA)-associated responses at later stages in the defence response. AtWRKY40 is a transcriptional repressor regulating abscisic acid and abiotic stress responses (Chen *et al*. 2010). Recently, Pandey *et al*. (2010) demonstrated that WRKY40 negatively modulates the expression of positive regulators of defence such as *EDS1* and *PAD4*, but positively modulates the expression of some key JA-signalling genes by partially suppressing the expression of the JAZ repressor. Similarly, AtWRKY75 has been shown to have roles in both basal and *R*-gene mediated defence responses (Encinas-Villarejo *et al*. 2009). However, AtMYB111 and AtWRKY31 have not been previously associated with defence. The function of AtWRKY31 is unknown (The *Arabidopsis* Information Resource, www.arabidopsis.org, accessed December 2012) and AtMYB111 is known to regulate the flavanol branch of flavonoid biosynthesis in various organs (Stracke *et al*. 2010). Although the soybean homologue *AtWRKY31* is expressed at relatively low levels, the soybean homologue *AtMYB111* is strongly expressed across all most time-points. Both of these genes would be key targets for future functional analyses (Fig. 4*a*).

Our analyses also identified transcription factors and receptors uniquely differentially expressed only in resistant reactions including homologues of *AtL6*, *AtMPK3*, *AtSZF2*, *AtWRKY11*, *AtWRKY33*, *AtWRKY40*, *AtWRKY51*, *AtWRKY53*, *AtZAT10* and *AtCRK9*. The roles of *AtWRKY33* and *AtWRKY40* are discussed above, but it is important to note that the soybean genes differentially expressed only in the resistant reactions are homologues of those expressed across all reactions. Similarly, the soybean *AtCRK9* homologue common to resistant and susceptible reactions is different from the homologue unique to resistance. Although AtWRKY11 is a negative regulator of basal defence responses (Journot-Catalino *et al*. 2006), AtWRKY53 is a positive regulator (Hu *et al*. 2012). AtWRKY51 is a negative regulator of JA-inducible defence responses (Gao *et al*. 2011). AtSZF2 and AtZAT10 are both regulators of abiotic stress responses (Mittler *et al*. 2006; Sun *et al*. 2007). The *Arabidopsis AtL6* gene encodes a RING-type ubiquitin ligase that controls defence and nitrogen/carbon responses (Maekawa *et al*. 2012) and is upregulated by the elicitors flg22 and chitin and by infection with *Pseudomonas syringae* pv. tomato DC3000. AtMAPK3 also plays a role in resistant responses to flg22 (Galletti *et al*. 2011). *AtCRK29* encodes a cysteine-rich receptor kinase with no known function. We note that the soybean *AtCRK29* homologue common to resistant and susceptible reactions was broadly expressed (Fig. 4*a*) whereas the homologue unique to resistant reactions was expressed mostly at earlier time points.

In addition to comparing gene expression across time course experiments, we used VIGS to identify genes and transcription factors essential to *Rpp4*-mediated signalling. Comparisons of the VIGS microarray dataset with the *Rpp2*, *Rpp3* and *Rpp4* time course experiments identified 72 genes that were unique to the VIGS experiment. Since this experiment was conducted at the latest time point after *P. pachyrhizi* inoculation (14 dai), it suggests many downstream defence genes have yet to be discovered. These genes may function to maintain defence response after the threat of *P. pachyrhizi* attack has passed. Within this dataset, we identified four transcription factors including soybean homologues of *AtBZIP9*, *AtTTL1*, *AtIAA17 and AtIDD2*. While roles in defence have not been demonstrated for *AtBZIP9* and *AtIDD2*, *AtTTL1* is required for osmotic stress tolerance (Lakhssassi *et al*. 2012) and *AtIAA17* increases tolerance to *Plasmodiophora brassicae* (Kazan and Manners 2009).

In conclusion, conducting microarray analysis over a 12 day time course demonstrated that *Rpp4*-mediated resistance occurs through the action of two different biphasic responses. Genes initially repressed by *P. pachyrhizi* infection have a three time point lag in expression while induced genes have a single time point lag. Initial gene expression at 12 hai is associated with basal defences, whereas later defence responses are likely *R*-gene mediated. Comparison of this dataset with the *Rpp2* and *Rpp3* microarray data developed by our group (van de Mortel *et al*. 2007; Schneider *et al*. 2011) revealed surprising differences. Of the probe sets present on the array, only 214 were common to all resistant and susceptible interactions. Further, only 54 probe sets were common only to resistance reactions. In each of these cases, only small numbers of transcription factors regulate complex defence networks. By

combining VIGS with microarray analysis, we could identify defence genes directly responding to the action of a specific *R*-gene. Furthermore, we could use bioinformatics approaches to identify transcription factors that potentially regulate the expression of downstream defence genes. Combined, these approaches have revealed several interesting target genes to test for their roles in soybean responses to *P. pachyrhizi* and other pathogens.

Acknowledgements

The authors are grateful for financial support provided by the United Soybean Board, the North Central Soybean Research Program, the Iowa Soybean Association, the National Science Foundation Plant Genome Research Program (award number 0820642), the National Council for Scientific and Technological Development (CNPq, Brazil) and the US Department of Agriculture-Agricultural Research Service. EEO/Non-Discrimination Statement: The USDA is an equal opportunity provider and employer. Disclaimer: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

References

Agarwal PK, Jha B (2010) Transcription factors in plants and ABA dependent and independent abiotic stress signaling. *Biologia Plantarum* **54**, 201–212. doi:10.1007/s10535-010-0038-7

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402. doi:10.1093/nar/25.17.3389

Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh L-SL (2004) UniProt – the universal protein knowledgebase. *Nucleic Acids Research* **32**, D115–D119. doi:10.1093/nar/gkh131

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT , *et al*. (2000) Gene ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29. doi:10.1038/75556

Assunção AGL, Herrero E, Lin Y, Huettel B, Talukdar S, Smaczniak C, Immink RGH, van Eldik M, Fiers M, Schat H, Aarts MGM (2010) *Arabidopsis thaliana* transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 10 296–10 301. doi:10.1073/pnas.1004788107

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1994) 'Current protocols in molecular biology.' (John Wiley & Sons: Hoboken, NJ)

Birkenbihl RP, Diezel C, Somssich IE (2012) Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses toward *Botrytis cinerea* infection. *Plant Physiology* **159**, 266–285. doi:10.1104/pp.111.192641

Bolton MD (2009) Primary metabolism and plant defense: fuel for the fire. *Molecular Plant-Microbe Interactions* **22**, 487–497. doi:10.1094/MPMI-22-5-0487

Bonde MR, Nester SE, Austin CN, Stone CL, Frederick RD, Hartman GL, Miles MR (2006) Evaluation of virulence of *Phakopsora pachyrhizi* and *P. meibomiae* isolates. *Plant Disease* **90**, 708–716. doi:10.1094/PD-90-0708

Bonferroni CE (1935) Ill Calcolo delle assicurazioni su gruppi di teste. In 'Studi in Onore del Professore Salvatore Ortu Carboni'. pp. 13–60. (Rome, Italy)

Bromfield KR, Hartwig EE (1980) Resistance to soybean rust and mode of inheritance. *Crop Science* **20**, 254–255.

Chakraborty N, Curley J, Frederick RD, Hyten DL, Nelson RL, Hartman GL, Diers B (2009) Mapping and confirmation of a new allele at *Rpp1* from soybean PI 594538A conferring RB lesion-type resistance to soybean rust. *Crop Science* **49**, 783–790. doi:10.2135/cropsci2008.06.0335

Chen H, Lai Z, Shi J, Xiao Y, Chen Z, Xu X (2010) Roles of *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biology* **10**, 281. doi:10.1186/1471-2229-10-281

Chen L, Song Y, Li S, Zhang L, Zou C, Yu D (2012) The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta* **1819**, 120–128. doi:10.1016/j.bbagrm.2011.09.002

Cheong YH, Chang HS, Gupta R, Wang X, Zhu T, Luan S (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. *Plant Physiology* **129**, 661–677. doi:10.1104/pp.002857

Choi JJ, Alkarouf NW, Schneider KT, Matthews BF, Frederick RD (2008) Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. *Functional & Integrative Genomics* **8**, 341–359. doi:10.1007/s10142-008-0080-0

de Hoon MJ, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. *Bioinformatics* **20**, 1453–1454. doi:10.1093/bioinformatics/bth078

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in *Arabidopsis*. *Trends in Plant Science* **15**, 573–581. doi:10.1016/j.tplants.2010.06.005

Encinas-Villarejo S, Maldonado AM, Amil-Ruiz F, de los Santos B, Romero F, Pliego-Alfaro F, Munoz-Blanco J, Caballero JL (2009) Evidence for a positive regulatory role of strawberry (*Fragaria* × *ananassa*) Fa WRKY1 and *Arabidopsis* At WRKY75 proteins in resistance. *Journal of Experimental Botany* **60**, 3043–3065. doi:10.1093/jxb/erp152

Fisher RA (1966) 'The design of experiments.' 8th edn. (Oliver and Boyd: Edinburgh)

Frith MC, Fu Y, Chen JF, Hansen U, Weng Z (2004) Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Research* **32**, 1372–1381. doi:10.1093/nar/gkh299

Galletti R, Ferrari S, De Lorenzo G (2011) *Arabidopsis* MPK3 and MPK6 play different roles in basal and oligogalacturonide- or flagellin-induced resistance against *Botrytis cinerea*. *Plant Physiology* **157**, 804–814. doi:10.1104/pp.111.174003

Gao QM, Venugopal S, Navarre D, Kachroo A (2011) Low oleic acid-derived repression of jasmonic acid-inducible defense responses requires the WRKY50 and WRKY51 proteins. *Plant Physiology* **155**, 464–476. doi:10.1104/pp.110.166876

Garcia A, Calvo ES, de Souza Kiihl R, Harada A, Hiromoto DM, Vieira LG (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. *Theoretical and Applied Genetics* **117**, 545–553. doi:10.1007/s00122-008-0798-z

Giraud E, Aken O, Ho LH, Whelan J (2009) The transcription factor ABI4 is a regulator of mitochondrial retrograde expression of *ALTERNATIVE OXIDASE1a*. *Plant Physiology* **150**, 1286–1296. doi:10.1104/pp.109.139782

Graham MA (2012) Identification of genes that mediate protection against soybean pathogens. In 'Designing soybeans for the 21st century'. (Ed. R Wilson) pp. 97–108. (American Oil Chemists Society Press: Urbana, IL)

Grotewold E, Drummond BJ, Bowen B, Peterson T (1994) The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell* **76**, 543–553. doi:10.1016/0092-8674(94)90117-1

Hartwig EE (1986) Identification of a fourth major gene conferring resistance to soybean rust. *Crop Science* **26**, 1135–1136. doi:10.2135/cropsci1986.0011183X002600060010x

Hartwig EE, Bromfield KR (1983) Relationships among 3 genes conferring specific resistance to rust in soybeans. *Crop Science* **23**, 237–239. doi:10.2135/cropsci1983.0011183X002300020012x

Hu Y, Dong Q, Yu D (2012) *Arabidopsis* WRKY46 coordinates with WRKY70 and WRKY53 in basal resistance against pathogen *Pseudomonas syringae. Plant Science* **185-186**, 288–297. doi:10.1016/j.plantsci.2011.12.003

Hückelhoven R (2007) Cell wall-associated mechanisms of disease resistance and susceptibility. *Annual Review of Phytopathology* **45**, 101–127. doi:10.1146/annurev.phyto.45.062806.094325

Hyten DL, Hartman GL, Nelson RL, Frederick RD, Concibido VC, Narvel JM, Cregan PB (2007) Map location of the *Rpp1* locus that confers resistance to soybean rust in soybean. *Crop Science* **47**, 837–840. doi:10.2135/cropsci2006.07.0484

Hyten DL, Smith JR, Frederick RD, Tuchekr ML, Song Q, Cregan PB (2009) Bulked segregant analysis using the GoldenGate assay to locate the *Rpp*3 locus that confers resistance to soybean rust in soybean. *Crop Science* **49**, 265–271. doi:10.2135/cropsci2008.08.0511

Journot-Catalino N, Somssich IE, Roby D, Kroi T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana. The Plant Cell* **18**, 3289–3302. doi:10.1105/tpc.106.044149

Kang JY, Choi HI, Im MY, Kim SY (2002) Arabidopsis basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *The Plant Cell* **14**, 343–357. doi:10.1105/tpc.010362

Kawaoka A, Ebinuma H (2001) Transcriptional control of lignin biosynthesis by tobacco LIM protein. *Phytochemistry* **57**, 1149–1157. doi:10.1016/S0031-9422(01)00054-1

Kazan K, Manners JM (2009) Linking development to defense: auxin in plant-pathogen interactions. *Trends in Plant Science* **14**, 373–382. doi:10.1016/j.tplants.2009.04.005

Kim KS, Unfried JR, Hyten DL, Frederick RD, Hartman GL, Nelson RL, Song Q, Diers BW (2012) Molecular mapping of soybean rust resistance in soybean accession PI561356 and SNP haplotype analysis of the *Rpp1* region in diverse germplasm. *Theoretical and Applied Genetics* **125**, 1339–1352. doi:10.1007/s00122-012-1932-5

Kumudini S, Godoy CV, Board JE, Omielan J, Tollenaar M (2008) Mechanisms involved in soybean rust-induced yield reduction. *Crop Science* **48**, 2334–2342. doi:10.2135/cropsci2008.01.0009

Lakhssassi N, Doblas VG, Rosado A, Estaban del Valle A, Posé D, Jimenez AJ, Castillo AG, Valpuesta V, Borsani O, Botella MA (2012) The *Arabidopsis thaliana TETRATRICO PEPTIDE THIOREDOXIN-LIKE* gene family is required for osmotic stress tolerance and male sporogenesis. *Plant Physiology* **158**, 1252–1266. doi:10.1104/pp.111.188920

Li S, Smith JR, Ray JD, Frederick RD (2012) Identification of a new soybean rust resistance gene in PI 567102B. *Theoretical and Applied Genetics* **125**, 133–142. doi:10.1007/s00122-012-1821-y

Lindermayr C, Sell S, Muller B, Leister D, Durner J (2010) Redox regulation of the NPR1–TGA1 system of *Arabidopsis thaliana* by nitric oxide. *The Plant Cell* **22**, 2894–2907. doi:10.1105/tpc.109.066464

Liu JZ, Horstman HD, Braun E, Graham MA, Zhang C, Navarre D, Qiu Q, Lee L, Nettleton D, Hill J, Whitham SA (2011) Soybean homologs of *MPK4* negatively regulate defense responses and positively regulate growth and development. *Plant Physiology* **157**, 1363–1378. doi:10.1104/pp.111.185686

Maekawa S, Sato T, Asada Y, Yasuda S, Yoshida M, Chiba Y, Yamaguchi J (2012) The *Arabidopsis* ubiquitin ligases ATL31 and ATL6 control the defense response as well as the carbon/nitrogen response. *Plant Molecular Biology* **79**, 217–227. doi:10.1007/s11103-012-9907-0

Major IT, Nicole MC, Duplessis S, Seguin A (2010) Photosynthetic and respiratory changes in leaves of poplar elicited by rust infection. *Photosynthesis Research* **104**, 41–48. doi:10.1007/s11120-009-9507-2

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E (2006) TRANSFAC® and its module TRANSCompel®: transcriptional gene regulation in eukaryotes. *Nucleic Acids Research* **34**, D108–D110. doi:10.1093/nar/gkj143

McLean RJ, Byth D (1980) Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybean. *Australian Journal of Agricultural Research* **31**, 951–956. doi:10.1071/AR9800951

Melching JS, Bromfield KR, Kingsolver CH (1983) The plant pathogen containment facility at Frederick, Maryland. *Plant Disease* **67**, 717–722. doi:10.1094/PD-67-717

Meyer JDF, Silva DCG, Yang C, Pedley KF, Zhang C, van de Mortel M, Hill JH, Shoemaker RC, Abdelnoor RV, Whitham SA, Graham MA (2009) Identification and analyses of candidate genes for *Rpp*4-mediated resistance to Asian soybean rust in Soybean. *Plant Physiology* **150**, 295–307. doi:10.1104/pp.108.134551

Mittler R, Kim Y, Song L, Coutu J, Coutu A, Ciftci-Yilmaz S, Lee H, Stevenson B, Zhu JK (2006) Gain- and loss-of-function mutations in *Zat10* enhance the tolerance of plants to abiotic stress. *FEBS Letters* **580**, 6537–6542. doi:10.1016/j.febslet.2006.11.002

Monteros MJ, Missaoui AM, Phillips DV, Walker DR, Boerma HR (2007) Mapping and confirmation of the 'Hyuuga' red-brown lesion resistance gene for Asian soybean rust. *Crop Science* **47**, 829–834. doi:10.2135/cropsci06.07.0462

Mukherjee K, Choudhury AR, Gupta B, Gupta S, Sengupta DN (2006) An ABRE-binding factor, OSBZ8, is highly expressed in salt tolerant cultivars than in salt sensitive cultivars of indica rice. *BMC Plant Biology* **6**, 18. doi:10.1186/1471-2229-6-18

Nettleton D (2006) A discussion of statistical methods for design and analysis of microarray experiments for plant scientists. *The Plant Cell* **18**, 2112–2121. doi:10.1105/tpc.106.041616

Ohme-Takagi KS, Shinshi H (2000) Regulation of ethylene-induced transcription of defense genes. *Plant & Cell Physiology* **41**, 1187–1192. doi:10.1093/pcp/pcd057

Pandey SP, Roccaro M, Schon M, Logemann E, Somssich IE (2010) Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of *Arabidopsis*. *The Plant Journal* **64**, 912–923. doi:10.1111/j.1365-313X.2010.04387.x

Pandey AK, Yang C, Zhang C, Graham MA, Horstman HD, Lee Y, Zabotina OA, Hill JH, Pedley KF, Whitham SA (2011) Functional analysis of the Asian soybean rust resistance pathway mediated by *Rpp2. Molecular Plant-Microbe Interactions* **24**, 194–206. doi:10.1094/MPMI-08-10-0187

Panthee DR, Yuan JS, Wright DL, Marois JJ, Mailhot D, Stewart CN Jr (2007) Gene expression analysis in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in an early growth stage. *Functional & Integrative Genomics* **7**, 291–301. doi:10.1007/s10142-007-0045-8

Panthee DR, Marois JJ, Wright DL, Narva'ez D, Yuan JS, Stewart CN Jr (2009) Differential expression of genes in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) is soybean growth stage-specific. *Theoretical and Applied Genetics* **118**, 359–370. doi:10.1007/s00122-008-0905-1

R Development Core Team (2006) 'R: A language and environment for statistical computing.' (R Foundation for Statistical Computing: Vienna, Austria)

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. *Trends in Plant Science* **15**, 247–258. doi:10.1016/j.tplants.2010.02.006

Saldanha AJ (2004) Java TreeView: extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248. doi:10.1093/bioinformatics/bth349

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178–183. doi:10.1038/nature08670

Schneider KT, van de Mortel M, Bancroft TJ, Braun E, Nettleton D, Nelson RT, Frederick RD, Baum TJ, Graham MA, Whitham SA (2011) Biphasic gene expression changes elicited by *Phakopsora pachyrhizi* in soybean correlate with fungal penetration and haustoria formation. *Plant Physiology* **157**, 355–371. doi:10.1104/pp.111.181149

Sell S, Hehl R (2004) Functional dissection of a small anaerobically induced bZIP transcription factor from tomato. *European Journal of Biochemistry* **271**, 4534–4544. doi:10.1111/j.1432-1033.2004.04413.x

Silva DCG, Yamanaka N, Brogin RL, Arias CAA, Nepomuceno AL, Di Mauro AO, Pereira SS, Nogueira LM, Passianotto ALL, Abdelnoor RV (2008) Molecular mapping of two loci that confer resistance to Asian rust in soybean. *Theoretical and Applied Genetics* **117**, 57–63. doi:10.1007/s00122-008-0752-0

Sinclair JB (1989) Threats to soybean production in the tropics: red leaf blotch and leaf rust. *Plant Disease* **73**, 604–606. doi:10.1094/PD-73-0604

Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* **3**, article 3. doi:10.2202/1544-6115.1027

Smyth GK (2005) Limma: linear models for microarray data. In 'Bioinformatics and computational biology solutions using R and Bioconductor'. (Eds R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber) pp. 397–420. (Springer: New York)

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9440–9445. doi:10.1073/pnas.1530509100

Stracke R, Favory JJ, Gruber H, Bartelniewoehner L, Bartels S, Binkert M, Funk M, Weissha B, Ulm R (2010) The *Arabidopsis* bZIP transcription factor HY5 regulates expression of the *PFG1/MYB12* gene in response to light and ultraviolet-B radiation. *Plant, Cell & Environment* **33**, 88–103. doi:10.1111/j.1365-3040.2009.02061.x

Sugano S, Jianng C, Miyazawa S, Masumoto C, Yazawa K, Hayashi N, Shimono M, Nakayama A, Miyao M, Takatsuji H (2010) Roles of OsNPR1 in rice defense program as revealed by genome-wide expression analysis. *Plant Molecular Biology* **74**, 549–562. doi:10.1007/s11103-010-9695-3

Sun J, Liang H, Xy Y, Li H, Wy X, Xie Q, Li C (2007) The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in *Arabidopsis*. *Plant & Cell Physiology* **48**, 1148–1158. doi:10.1093/pcp/pcm088

Tremblay A, Hosseini P, Alkharouf NW, Li S, Matthews BF (2010) Transcriptome analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. *Plant Science* **179**, 183–193. doi:10.1016/j.plantsci.2010.04.011

Tremblay A, Hosseini P, Li S, Alkarouf NW, Matthews BF (2012) Identification of genes expressed by *Phakopsora pachyrhizi*, the pathogen causing soybean rust, at a late stage of infection of susceptible soybean leaves. *Plant Pathology* **61**, 773–786. doi:10.1111/j.1365-3059.2011.02550.x

van de Mortel M, Recknor JC, Graham MA, Nettleton D, Dittman JD, Nelson RT, Godoy CV, Abdelnoor RV, Almeida AMR, Baum TJ, Whitham SA (2007) Distinct biphasic mRNA changes in response to Asian soybean rust infection. *Molecular Plant-Microbe Interactions* **20**, 887–899. doi:10.1094/MPMI-20-8-0887

Voegele RT, Mendgen K (2003) Rust haustoria: nutrient uptake and beyond. *New Phytologist* **159**, 93–100. doi:10.1046/j.1469-8137.2003.00761.x

Wang A, Libaut M, Joshi T, Valliyodan B, Nguyen HT, Xu D, Stacey G, Cheng J (2010) SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biology* **10**, 14. doi:10.1186/1471-2229-10-14

Wei W, Huang J, Hao Y, Zou H, Wang H, Zhao J, Liu X, Zhang W, Ma B, Zhang JS, Chen SY (2009) Soybean GmPHD-type transcription regulators improve stress tolerance in transgenic Arabidopsis plants. *PLoS ONE* **4**, e7209. doi:10.1371/journal.pone.0007209

Wingender E, Dietze P, Karas H, Knüppel R (1996) TRANSFAC®: a database on transcription factors and their DNA binding sites. *Nucleic Acids Research* **24**, 238–241. doi:10.1093/nar/24.1.238

Wise RP, Caldo RA, Hong L, Shen L, Cannon EK, Dickerson JA (2007) BarleyBase/PLEXdb: a unified expression profiling database for plants and plant pathogens. *Methods in Molecular Biology (Clifton, N.J.)* **406**, 347–363.

Xue GP, Way HM, Richardson T, Drenth JM, Joyce PA, McIntyr CL (2011) Overexpression of *TaNAC69* leads to enhanced transcript levels of stress up-regulated genes and dehydration tolerance in bread wheat. *Molecular Plant* **4**, 697–712. doi:10.1093/mp/ssr013

Yamaoka Y, Fujiwara Y, Kakishima M, Katsuya K, Yamada K, Hagiwara H (2002) Pathogenic races of *Phakopsora pachyrhizi* on soybean and wild host plants collected in Japan. *Journal of General Plant Pathology* **68**, 52–56. doi:10.1007/PL00013053

Yang S, Sweetman JP, Amirsadeghi S, Barghchi M, Huttly AK, Chung W, Twell D (2001) Novel anther-specific *myb(MYB?)* genes from tobacco as putative regulators of phenylalanine ammonia-lyase expression. *Plant Physiology* **126**, 1738–1753. doi:10.1104/pp.126.4.1738

Yorinori JT, Paiva WM, Frederick RD, Costamilan LM, Bertagnolli PF, Hartman GE, Godoy CV, Nunes J Jr (2005) Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Disease* **89**, 675–677. doi:10.1094/PD-89-0675