2016

Identification of an orthologous clade of peroxidases that respond to feeding by greenbugs (*Schizaphis graminum*) in C_4 grasses

Erin D. Scully
USDA-ARS

Teresa Donze-Reiner
West Chester University

Haichuan Wang
University of Nebraska-Lincoln, hwang4@unl.edu

Thomas Eickhoff
University of Nebraska-Lincoln

Fred Baxendale
University of Nebraska--Lincoln, fbaxendale1@unl.edu

See next page for additional authors

CSIRO PUBLISHING

Functional Plant Biology, 2016, **43**, 1134–1148
http://dx.doi.org/10.1071/FP16104

Identification of an orthologous clade of peroxidases that respond to feeding by greenbugs (*Schizaphis graminum*) in C_4 grasses

Erin D. Scully[A], *Teresa Donze-Reiner*[B], *Haichuan Wang*[C], *Thomas E. Eickhoff*[C,D], *Frederick Baxendale*[C], *Paul Twigg*[E], *Frank Kovacs*[F], *Tiffany Heng-Moss*[C], *Scott E. Sattler*[G] *and Gautam Sarath*[G,H]

[A]Stored Product Insect and Engineering Research Unit, Center for Grain and Animal Health Research USDA-ARS, Manhattan, KS 66502, USA.
[B]Department of Biology, West Chester University, West Chester, PA 19382, USA.
[C]Department of Entomology, University of Nebraska-Lincoln, Lincoln, NE 68583, USA.
[D]Monsanto Company, 800 North Lindbergh Blvd, St Louis, MO 63167, USA.
[E]Department of Biology, University of Nebraska-Kearney, Kearney, NE 68849, USA.
[F]Department of Chemistry, University of Nebraska-Kearney, Kearney, NE 68849, USA.
[G]Grain, Forage and Bioenergy Research Unit, USDA-ARS, Lincoln, NE 68583, USA.
[H]Corresponding author. Email: gautam.sarath@ars.usda.gov

Abstract. Knowledge of specific peroxidases that respond to aphid herbivory is limited in C_4 grasses, but could provide targets for improving defence against these pests. A sorghum (*Sorghum bicolor* (L.) Moench) peroxidase (SbPrx-1; Sobic.002G416700) has been previously linked to biotic stress responses, and was the starting point for this study. Genomic analyses indicated that *SbPrx*-1 was part of a clade of five closely related peroxidase genes occurring within a ~30 kb region on chromosome 2 of the sorghum genome. Comparison of this ~30-kb region to syntenic regions in switchgrass (*Panicum virgatum* L.) and foxtail millet (*Setaria italica* L.) identified similar related clusters of peroxidases. Infestation of a susceptible sorghum cultivar with greenbugs (*Shizaphis graminum* Rondani) induced three of the five peroxidases. Greenbug infestation of switchgrass and foxtail millet plants showed similar inductions of peroxidases. SbPrx-1 was also induced in response to aphid herbivory in a greenbug-resistant sorghum line, Cargill 607E. These data indicate that this genomic region of C_4 grasses could be valuable as a marker to assess potential insect resistance in C_4 grasses.

Additional keywords: defense, defence, foxtail millet, greenbugs, plant resistance, sorghum, switchgrass, synteny.

Received 17 March 2016, accepted 29 July 2016, published online 26 August 2016

Introduction

Plants have developed elaborate defence systems to protect against aphid herbivory and other biotic stresses (Carmona and Fornoni 2013). One component of this aphid-defence system includes changes in the levels of stress-related proteins that can be directly or indirectly related to the cellular responses stimulated by insect feeding (Louis and Shah 2013). Therefore, documenting the induction of stress-related genes and proteins upon exposure to aphids can provide important information on the extent and severity of the plant stress response, as well as the ability of the plant to tolerate or overcome the negative effects of aphid herbivory.

Changes in cellular redox state and the accompanying increased production of reactive oxygen species (ROS) are often direct consequences of insect feeding (Moloi and van der Westhuizen 2008; Kerchev *et al*. 2012; Mai *et al*. 2013; Wu *et al*. 2013). How plants respond to elevated levels of ROS and oxidative stress may drive plant resistance and tolerance mechanisms to hemipteran pests (Gutsche *et al*. 2009*a*). In plant cells, increased activity of antioxidant enzymes, such as catalases, peroxidases and superoxide dismutases, generally accompanies insect feeding (Heng-Moss *et al*. 2004; Dowd and Lagrimini 2006; Franzen *et al*. 2007; Gulsen *et al*. 2010; Mai *et al*. 2013). These enzymes presumably act to restore redox balance and protect cells from oxidative damage while maintaining adequate levels of ROS that can interfere with insect feeding. This delicate balancing act could readily reach a tipping point, if insect pressure becomes insurmountable or if plants are unable to sustain the defensive responses against ROS to prevent cellular damage (Gutsche *et al*. 2009*b*). Among the key components of plant cellular defenses are the class III peroxidases. These heme-containing enzymes can both

produce and consume hydrogen peroxide (H_2O_2), and are strongly upregulated in plants exposed to biotic and abiotic stress. The genomes of most plants contain in excess of 100 class III peroxidases that can be classified into several different subgroups (Passardi *et al*. 2005; Saathoff *et al*. 2013). Individual genes belonging to distinct subgroups can be differentially regulated in response to stress and other environmental cues (Passardi *et al*. 2004*a*). Several studies (Dowd *et al*. 2006; Barbehenn *et al*. 2010; Suzuki *et al*. 2012) have demonstrated that plants engineered to overexpress peroxidases are more tolerant to specific insects, indicating a causal relationship between higher peroxidase activity and resistance and/or tolerance to insect herbivores. The identification of specific plant peroxidases involved in responding to insect herbivory, however, is lacking in many species.

The availability of multiple annotated grass genomes has permitted the identification of syntenic and microsyntenic regions across species (Massa *et al*. 2011; Schnable *et al*. 2012; Schnable 2015). Genes present in syntenic regions are often orthologous and may perform similar functions across species (Schnable 2015). Although a direct correspondence for the function and control of genes involved in primary metabolic pathways is often observed across species, the same may not hold true for other classes of genes. In expansive gene families comprised of large numbers of duplicated genes, it is possible that even the nearest orthologs in different species might not function in the same manner as neofunctionalisation, subfunctionalisation and other evolutionary changes are common after gene duplication events (Distelfeld *et al*. 2012). However, syntenic relationships may be useful for predicting functional outcomes, developing markers for resistance to biotic stresses and deciphering the regulatory circuits that control defence-related gene expression in the face of insect herbivory.

In a previous study, a single sorghum (*Sorghum bicolor* (L.) Moench) peroxidase (Sobic.002G146700, referred to as *SbPrx-1*) responded strongly after challenge with the pathogen *Colletotrichum*, the causative agent of anthracnose (Pratt *et al*. 2005; Zamora *et al*. 2009), and was part of a resistance response. In many plants, exposure to either pathogens or insects can elicit induction of the same defensive genes, although other aspects and outcomes of the defensive response can vary (Thaler *et al*. 2012; Lei *et al*. 2014). Peroxidases are also integral to the resistance/tolerance response in plants infested with piercing-sucking insects (Ni *et al*. 2001; Heng-Moss *et al*. 2004; Passardi *et al*. 2005; Gutsche *et al*. 2009*a*; Gulsen *et al*. 2010; Marchi-Werle *et al*. 2014). Sorghum–greenbug interactions have been well studied (Michels and Burd 2007) and aphid-resistant sorghum lines have been previously reported (Dogramaci *et al*. 2007). However, the identification of peroxidases functionally linked to defence responses in sorghum as well as orthologous peroxidases that respond to biotic stresses in multiple species of grasses have not been investigated. In this study, the role of a syntenic region containing orthologous groups of peroxidases in response to aphid feeding was investigated in the genomes of three closely related C_4 grass species, sorghum, switchgrass (*Panicum virgatum* L.) and foxtail millet (*Setaria italica* L.).

Materials and methods

Insects

Schizaphis graminum (greenbug; Biotype I) used for sorghum, switchgrass and foxtail millet infestations were maintained on BCK60 sorghum plants in a growth chamber at $25 \pm 1°C$, $40 \pm 10\%$ RH and a 1410 hours (L:D) photoperiod by the Department of Entomology at the University of Nebraska-Lincoln as described previously (Koch *et al*. 2014).

Phylogenetic and syntenic studies

To identify other sorghum peroxidase genes on chromosome 2 that may also respond to aphid feeding, BLASTP searches were performed using SbPrx-1 as a query. In all, a total of 25 related sequences classified as class III peroxidases were obtained and used for further phylogenetic analyses to determine their evolutionary relationships. Peroxidase amino acid sequences were aligned using ClustalW and neighbour-joining trees were constructed using Mega 5.0 (Kumar *et al*. 2008) using the following parameters: 1000 bootstrap pseudoreplicates, the JTT + G evolutionary model, homogeneous evolutionary rates among lineages, and gaps were treated as pairwise deletions. An approximate likelihood ratio (aLRT) test (Anisimova and Gascuel 2006) was used to assess the probability of the branching topology on the bootstrap consensus trees using the program PAML (Yang 2007). Further, to determine whether the clade of peroxidases (Clade 3) that was most closely related to SbPrx-1 shared a common evolutionary ancestry, a cladogram was generated using all of the sorghum peroxidases belonging to the SbPrx-1 clade and the putative ancestral rice (*Oryza sativa* L.), sorghum and liverwort (*Marchantia polymorpha* L.) sequences. The rice and liverwort sequences had been identified earlier (Passardi *et al*. 2004*a*) and the predicted sorghum orthologue of the ancestral rice and liverwort sequence was identified using reciprocal BLASTP searches.

Expression levels of Clade 3 sorghum peroxidases in response to aphid feeding

To determine whether the Clade 3 peroxidases responded to aphid feeding in sorghum, RT-qPCR was performed on infested and uninfested sorghum plants 10 days after aphid infestation (DAI). A single time point was selected for two reasons: (1) previous physiological studies on other species have indicated that induction of peroxidases was time dependent, with maximal enzyme activity detected between 7 and 15 DAI (Ni *et al*. 2001; Heng-Moss *et al*. 2004); and (2) visual observations during the routine maintenance of greenbugs on sorghum (BCK60) plants indicated intermediate damage such as yellowing and loss of vigour to plants around 10 DAI (Koch *et al*. 2014), suggesting that plant defence responses were likely well induced at this time point. However, it is possible that specific peroxidase encoding genes could have different expression profiles, and one harvest time would not provide this data. Nevertheless, other studies that have assayed for peroxidase activity or expression have detected elevated peroxidase activity anywhere from 6 to 21 days in plants infested with shoot fly (*Atherigona soccata*) (Padmaja *et al*. 2014), Russian wheat aphid (*Diuraphis noxia*) (Gutsche *et al*.

2009*a*) and greenbugs (Chaman *et al*. 2001), suggesting that peroxidases that responded to aphid feeding could be readily detected at 10 DAI. However, future experiments, based on results presented here, can be designed to evaluate peroxidase and other plant defensive gene expression at both early and later stages using plants with well defined (resistant or susceptible) responses to greenbugs.

Plants were grown in 'SC 10 Super Cell' Single Cell conetainers to the three-leaf stage. Plants were infested with ~10 nymphs of *S. graminum* biotype I by placing them on the leaves using a fine eye brush as described previously (Koch *et al*. 2014). Infested and uninfested plants were caged individually and arranged in a randomised complete block design.

Leaves were collected at 10 DAI for RNA isolation. After removal of aphids with a camel hair brush, leaves were flash frozen with liquid nitrogen and stored at −80°C. Total RNA was extracted using TRIzol reagent (Invitrogen), purified using the RNeasy MinElute Cleanup Kit (Qiagen Inc.) and validated essentially as described by Scully *et al*. (2016). The total RNA samples were DNase treated and used for cDNA synthesis as previously described (Palmer *et al*. 2014) using SuperScript III reverse transcriptase (Invitrogen) and random primers according to the manufacturer's protocol. Relative expression levels of the Clade 3 peroxidase genes in were measured using qPCR with Bio-Rad SsoAdvanced SYBR Green supermix (Bio-Rad Laboratories) following the manufacturer's protocols: 20 µL qPCR reactions were performed in triplicate on each cDNA sample with 0.6 µM forward primer, 0.6 µM reverse primer (see Table S1, available as Supplementary Material to this paper), 1 µL 1 : 5 diluted cDNA and 10 µL SsoAdvanced SYBR Green Supermix (Bio-Rad) using the 7500 Fast Realtime PCR System (Applied Biosystems). Thermal cycling parameters were as follows: initial denaturation for 30 s at 95°C, 40 cycles of denaturation at 95°C for 5 s and annealing at 62°C for 10 s. Primer specificities were confirmed by dissociation curve analysis, consisting of denaturation at 95°C for 5 s, cooling to 65°C for 15 s and gradual heating at 0.2°C s^{-1} to a final temperature of 95°C. Relative expression was computed using the ΔΔCt method using the sorghum ubiquitin-conjugating enzyme (UCE) as a control (Table S1). No-template and no-RT controls were also analysed to verify the absence of DNA contamination. Relative expression values were statistically analysed using ANOVA followed by Tukey's honestly significant difference (HSD) post hoc analysis to determine which peroxidases were induced by aphid feeding in comparison to uninfested control plants. Primer sequences used for the peroxidase and control genes are shown in Table S1. Four biological replicates were analysed in triplicate.

Identification of orthologous Clade 3 peroxidases in switchgrass and foxtail millet

To determine whether the genomes of other C_4 grasses (e.g. switchgrass and foxtail millet) also harboured orthologs to the sorghum Clade 3 peroxidases, orthology searches were performed using Phytozome ver. 10.2 (Goodstein *et al*. 2012), which uses InParanoid to compute orthologue groups (Ostlund *et al*. 2010). Putative switchgrass, sorghum and foxtail millet orthologs of the sorghum Clade 3 peroxidases were subjected to maximum-likelihood based phylogenetic analysis using Garli (Zwicki 2006). WAG+I+G was chosen as the optimal evolutionary model based on AIC criteria using ProtTest (Abascal *et al*. 2005) and evolution was simulated for 500 000 generations or until likelihood scores reached convergence: 500 bootstrap pseudoreplications were performed. Synteny was assessed by comparing the order of the flanking genes in the genomes of the three plant species and assessing their orthology by recipricol BLASTP searches.

Response of Clade 3 peroxidase orthologs to aphid feeding in switchgrass and foxtail millet

Switchgrass (cv. Summer) and foxtail millet plants (SI-7 forage variety) were grown under the same conditions described for sorghum. Plants were grown to the three leaf stage and 2–3 aphid adults were transferred to plants using a fine hair paint brush. RNA was isolated from infested and uninfested control plants at 10 DAI and qPCR was performed using the same protocol described for sorghum. The actin2 (Act2) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) genes were used for normalisation in foxtail millet and switchgrass respectively. Primer sequences are listed in Table S1, and qPCR was performed on four biological replicates in triplicate and expression levels were computed using the ΔΔCt method with the respective control genes. Peroxidases whose expression levels were significantly impacted by aphid feeding compared with uninfested control plants were detected using ANOVA followed by Tukey's HSD post-hoc analysis.

Identification of cis*-regulatory elements in the promoters of peroxidases responsive to aphid feeding*

The 5′ putative promoter regions 2000 bp upstream from the ATG start codon of the Clade 3 peroxidases identified in sorghum, switchgrass and foxtail millet were retrieved from Phytozome and analysed for *cis*-acting regulatory DNA elements via a signal scan comparison to the PLACE database (http://www.dna.affrc.go.jp/PLACE/ accessed 15 November 2015) (Higo *et al*. 1999). NMDS analysis was conducted using the vegan library (Dixon 2003) in the R statistical environment (ver. 3.0 for Linux) to determine whether there were similar *cis* regulatory elements among the promoters of the aphid-responsive genes. The program DREME (Bailey 2011) was used to scan for novel motifs of 5–8 nucleotides in length enriched in the promoters of the aphid-responsive peroxidases relative to non-responsive peroxidase promoters.

Predictive structural modelling

I-Tas.SER (*I*terative *T*hreading *ASSE*mbly *R*efinement) (Zhang 2008) was used to perform predictive structural modelling of these proteins to attempt to gain structural insight and to determine whether there were any significant structural similarities of the peroxidases that responded to aphid feeding. The I-Tas.SER server uses an integrated set of programs for protein structure prediction. This modelling system typically works best when there is high sequence similarity between the proteins of interest and an experimentally determined structure or set of structures that already exists in the Protein Data Bank (PDB, found at www.rcsb.org, accessed 27 April 2015) (Berman

et al. 2003). Model reliability was assessed primarily using two structural measurements, the confidence score (C-score) and the TM-score (Zhang 2008). The C-score is based on threading template alignment quality and a set of structural assembly simulation convergence parameters. The TM-score is analogous to the RMSD measurement and is a measure of the structural similarity between two structures (Roy *et al*. 2010). Each of the amino acid sequences were uploaded to the I-Tas. SER server at http://zhanglab.ccmb.med.umich.edu/I-Tas.SER/, accessed 6 May 2015) for predictive modelling. All structure analysis and molecular image generation was performed using the UCSF Chimera package (Pettersen *et al*. 2004). Chimera is developed by the Resource for Biocomputing, Visualisation and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

Identification of Clade 3 peroxidase homologues in maize, rice and Brachypodium

To determine whether Clade 3 peroxidase orthologs were also present in the genomes of other C_4 grasses (*Zea mays* (maize)) and C_3 grasses (*Brachypodium distachyon* (*Brachypodium*) and rice), we used InParanoid to identify putative orthologous sequences in the genomes of these three plants. Maximum likelihood based phylogenies and assessments of synteny were performed as described above.

Identification of peroxidases that respond to aphid-feeding in a resistant sorghum cultivar (Cargill 607E and Garst 5715)

To determine whether any of these peroxidases could potentially be linked to greenbug resistant responses in sorghum, a resistant commercial sorghum hybrid Cargill 607E and another susceptible sorghum hybrid Garst 5715 were infested with aphids as described above and peroxidase activities were measured. Treatments were arranged in a $2 \times 2 \times 3$ factorial design consisting of two sorghum cultivars, two treatments (infested and non-infested plants) and three sampling dates (4, 8 and 12 DAI). Aphid numbers were counted and chlorophyll content of the first leaf was determined at each sampling date as described by Porra *et al*. (1989). Peroxidases were extracted from control and infested plants at 4, 10 and 12 DAI and analysed for activity and by gel electrophoresis at 12 DAI as described earlier (Heng-Moss *et al*. 2004).

A peroxidase band that was apparently constitutively expressed in both the uninfested resistant and susceptible plants (see 'Results'), but apparently upregulated in infested resistant plants was partially purified and identified as follows: soluble proteins were extracted by grinding ~7 g of tissue from Cargill 607E or Garst 5715 sorghum plants with dry ice followed by homogenisation with 20 mL of 20 mM HEPES (pH 7.2) containing 20 µL proteinase inhibitor cocktail and 1% (w/v) polyvinylpolypyrrolidone. The homogenate was passed through four layers of cheesecloth and centrifuged at 10 000*g* for 20 min at 4°C. Proteins in the supernatant were precipitated using ammonium sulfate (80% saturation). Precipitated proteins were redissolved in 2 mL of desalting buffer (25 mM Tris buffer, 1 mM $MgCl_2$, 1 mM $CaCl_2$ and 1 mM $MnCl_2$ pH 7.5 and 50 µL proteinase inhibitor cocktail) and centrifuged at 14 000*g* for 10 min at 4°C to pellet insolubles. The supernatant was desalted, equilibrated and eluted over a PD-10 desalting column (GE Healthcare Life Sciences) with desalting buffer. Approximately 2.5 mL of the desalted extract was passed over a Concanavalin A-Sepharose column (Con-A, Sigma) equilibrated with desalting buffer containing 0.5 M NaCl (wash buffer). The column was sequentially washed with 10 mL of wash buffer, followed by 10 mL wash buffer containing 0.5 M α-methyl mannopyranoside to elute bound proteins. Both the Con-A elutate and wash fractions were concentrated using YM-10 centricon filters as suggested by the manufacturer (Millipore Corp.). Concentrated protein samples were collected and 20 µL of sample was mixed with 5 µL loading buffer and separated by non-denaturing gel electrophoresis on 7.5% polyacrylamide gels (Heng-Moss *et al*. 2004). Zones of peroxidase activities were detected by soaking gels in 20 mL of 50 mM sodium acetate buffer (pH 5.0) containing 12 mg 4-hydroxy-3-methoxy phenyllacetic acid (dissolved in 0.5 mL of methanol) and 20 µL of 30% hydrogen peroxide for 10 min in 37°C water bath. Florescent bands of peroxidase activity were detected using a UV light source and one band specifically enriched in samples from 607E plants was excised using a razor blade. Excised gel slices were then subjected to automated trypsin digestion followed by mass spectrometry sequencing at the Mass Spectrometry Core Facility at the University of Nebraska-Lincoln (Kayser *et al*. 2006). Following sequencing, the identified peptide sequence was searched against standard databases. One peptide that matched exactly and mapped uniquely to SbPrx-1 was identified (Pratt *et al*. 2005).

Statistical analysis of data

Statistical analyses were performed in SAS (SAS Inc.) PROC MIXED ver. 9.2 or in Excel (Microsoft Corp.). Data analysis on RT-qPCR gene expression studies were completed according to manufacturer's protocol (Applied Biosystems). Relative expression values were statistically analysed using ANOVA followed by Tukey's honestly significant difference (HSD) post-hoc analysis to determine which peroxidases were induced by aphid feeding using $P \leq 0.05$ as a cutoff for significance.

Results

A clade of closely related class III peroxidases is located on chromosome 2 in sorghum

A BLASTP search of the sorghum genome using the Sobic.002G416700 (SbPrx-1) protein sequence as the query identified 25 additional sequences coding for other class III peroxidases on chromosome 2 and over 100 peroxidase-like sequences encoded throughout the rest of the sorghum genome (data not shown). A phylogenetic analysis of the 25 peroxidase sequences present on chromosome 2 indicated that there were three major clades containing two or more closely related peroxidases (Fig. 1*a*). The largest such clade (Clade 1) contained 11 genes (Fig. 1*a*). In general, peroxidases located in close proximity to one another on the chromosome were more

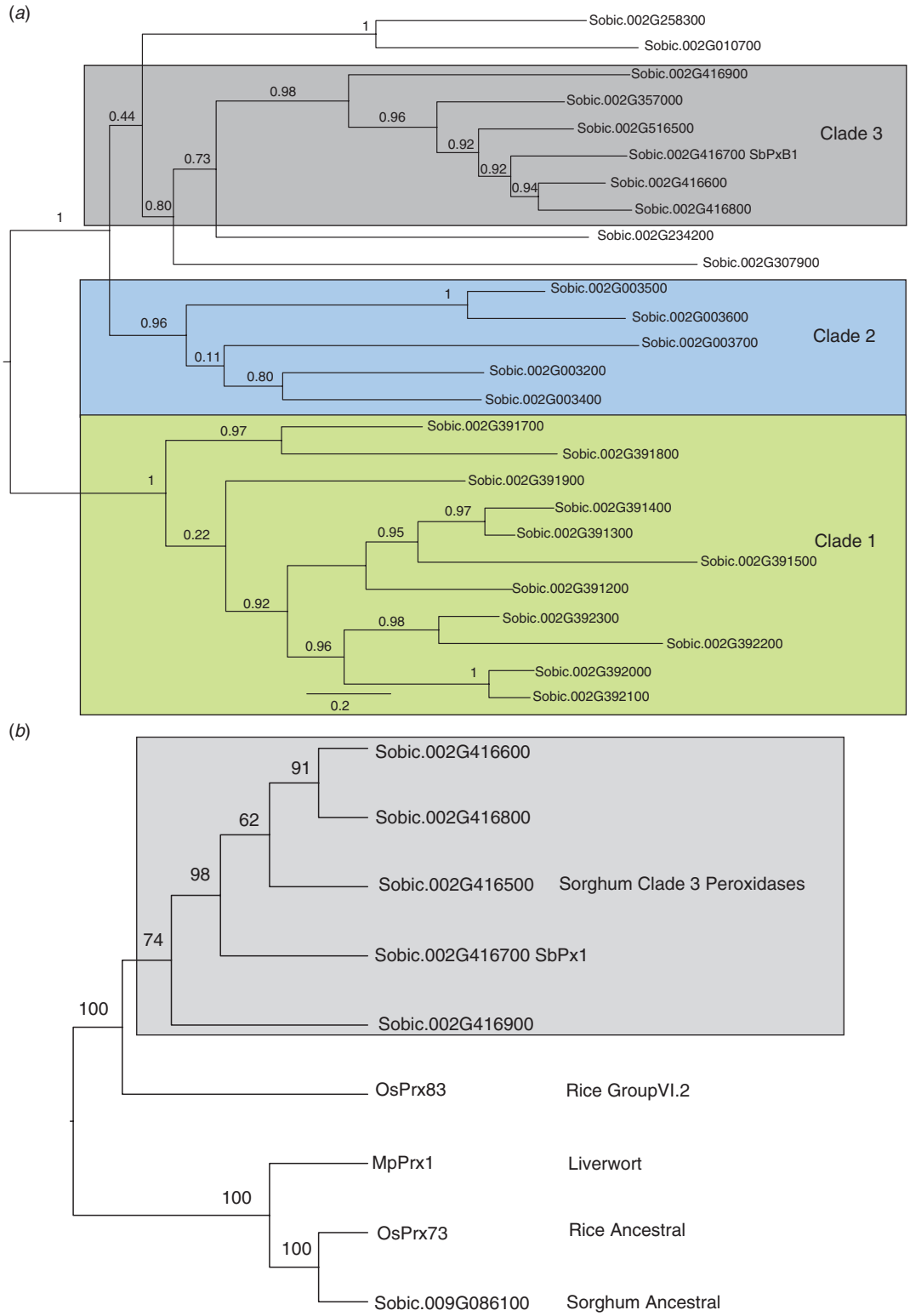


Fig. 1. Phylogenetic analyses of sorghum peroxidases. (*a*) Neighbour-joining analysis of sorghum class III/family 2 peroxidases localised to chromosome 2. Support for branching topology was generated using aLRT analysis. Values on the nodes represent the likelihood of each branch point. Scale bar represents number of substitutions per site. Based on this analysis, three clades of closely related peroxidases were apparent and are labelled Clades 1 through to 3. (*b*) Relationships of sorghum chromosome 2 clade 3 peroxidases to the putative ancestral rice and sorghum peroxidases. Neighbour-joining analysis was conducted using Mega 5. Bootstrap values on branches are shown as percentages.

closely related to each other than they were to peroxidases located along other regions of the chromosome (Fig. 1*a*). *SbPx-1* was found in a clade of six peroxidases (Clade 3), of which five were localised to a contiguous ~30 kb region of chromosome 2. These five peroxidases all share a common ancestor. This analysis indicated that Sobic.002G416900 was the most divergent of the peroxidases and could have served as the progenitor for the other peroxidases in this clade. In addition, the arrangement of these peroxidases along the chromosome is consistent with evolution via tandem duplication and/or gene amplification events.

Previously, an evolutionary grouping system formulated by Passardi *et al*. (2004*a*) was proposed with a selection of rice class III peroxidases. Based on reciprocal BLASTP searches, no rice orthologue to SbPx-1 could be conclusively identified; however, SbPx-1 had a highest scoring BLASTP alignment to *OsPrx-83*, which has been previously assigned to group IV.2, a group of monocot specific peroxidases (Passardi *et al*. 2004*a*) (Fig. 1*b*). Neighbour-joining analysis suggested that SbPx-1 and all of the class III sorghum peroxidases present in the ~30 kbp region on chromosome 2 belong to Group IV.2, providing further evidence for common evolutionary origin and potential functional similarities between these five sorghum peroxidases.

Three sorghum Clade 3 peroxidases are upregulated under aphid pressure

To investigate the responsiveness of the other Clade 3 sorghum peroxidases to aphid feeding, RT-qPCR analysis of the five peroxidases was conducted 10 DAI in the susceptible sorghum cultivar BCK60 (Fig. 2*a*). This analysis revealed that three of the five peroxidases were significantly induced ($P \leq 0.05$) upon aphid feeding, including Sobic.002G416500, Sobic.002G416600 and Sobic.002G416700. Sobic.002G416500 and Sobic.002G416700 (*SbPx-1*) were induced by 7.0- and 6.0-fold relative to uninfested control plants at 10 DAI ($P \leq 0.05$). Sobic.002G416600 was the most highly induced peroxidase in the clade and was upregulated by over 45-fold ($P \leq 0.05$) compared with uninfested control plants (Fig. 2*a*). In contrast, expression levels of Sobic.002G416800 and Sobic.002G416900 were similar in both the infested and uninfested plants (Fig. 2*a*).

Peroxidases are highly conserved within syntenic genomic regions in C_4 grasses

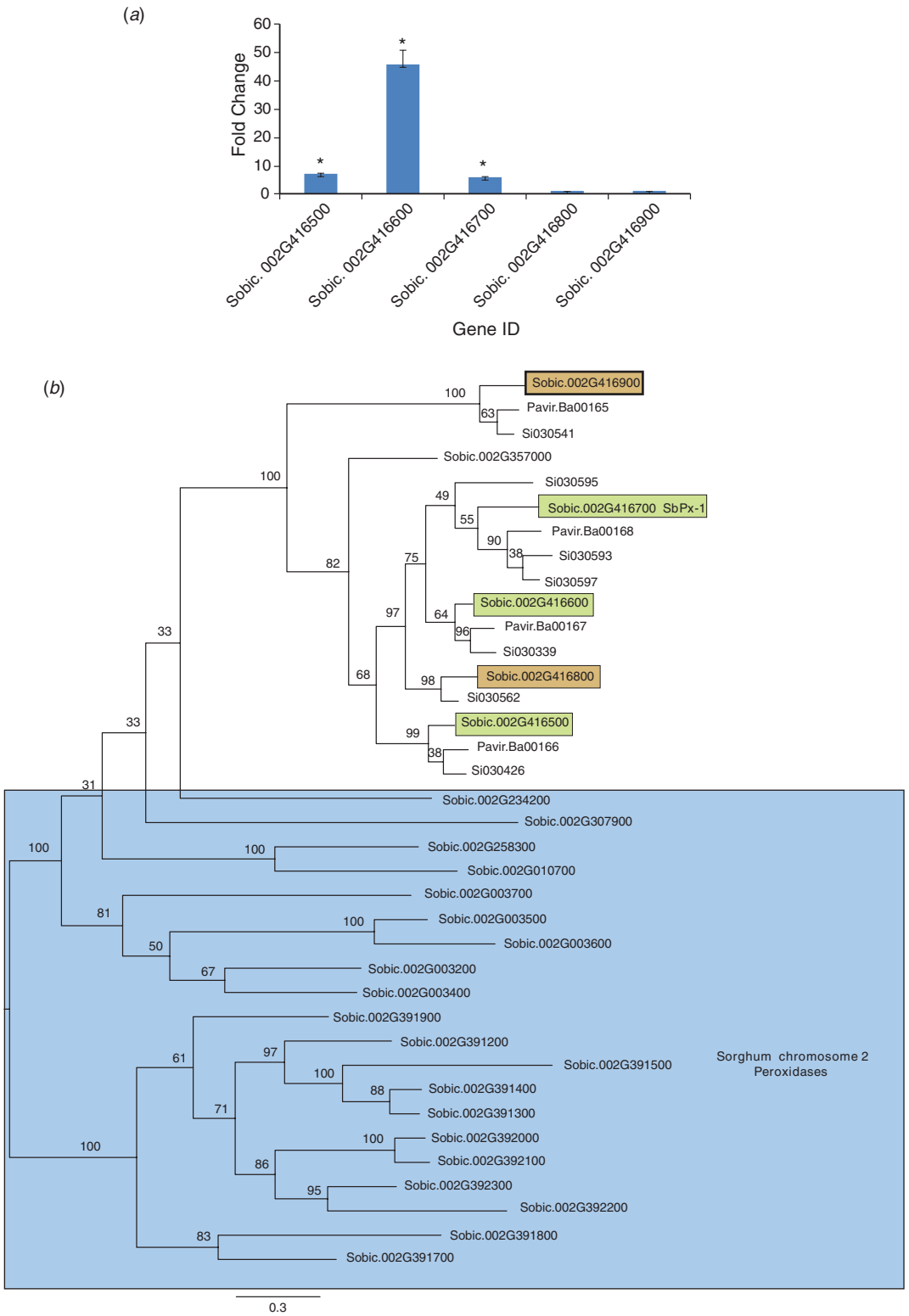
In general, the occurrence of clusters of phylogenetically related peroxidases throughout the genome appears to be specific to monocots as similar occurrences have not been observed in dicotyledonous species (i.e. *Arabidopsis*, *Populus*) (Passardi *et al*. 2004*a*). The recent release of the foxtail millet genome (Bennetzen *et al*. 2012) and ver. 1.1 of the switchgrass genome (www.phytozome.org, accessed 3 April 2015) permits a comparison of the syntenic genomic regions of these grasses. It also allows us to evaluate if peroxidases orthologous to Sobic.002G416700 (*SbPrx-1*) and other sorghum Clade 3 peroxidases contribute to response mechanisms to aphid feeding in other grasses.

Orthologous Clade 3 peroxidases were retrieved from the foxtail millet and switchgrass genomes using the Phytozome gene ancestry tool (Fig. 2*b*). The closest orthologs to *SbPrx-1* were Si030597 and PviBa00168 in foxtail millet and switchgrass respectively. In both cases, the orthologs to *SbPx-1* in the foxtail millet and switchgrass genomes occurred in regions of the genome containing multiple copies of class III peroxidases (see Fig. S1, available as Supplementary Material to this paper). In switchgrass, Pavir.Ba00168 occurred in an ~13 kb genomic region containing four peroxidases named Pavir.Ba00165 through Pavir.Ba00168 localised to a genomic scaffold assigned to chromosome 2a, whereas Si030597 was localised to a genomic interval of ~29 kb assigned to chromosome 2 containing seven peroxidases, including the gene models Si030426, Si030339, Si030593, Si030597, Si030562, Si030595 and Si030541. As in the cluster of peroxidases identified in sorghum, the switchgrass and foxtail millet peroxidases occurring within these clusters tended to be more closely related to each other than they were to other peroxidases occurring elsewhere on the chromosome (Figs S2, S3). The two exceptions to this observation were Si030541 (Fig. S2) and Pavir.Ba00165 (Fig. S3), which appeared to be more distantly related to the other peroxidases and instead, formed strongly supported clades with Si036729 and Pavir.Bb03620 respectively. Si036729 occurred as a singleton peroxidase on an unplaced genomic scaffold (scaffold_9:46953962.46955528 reverse) (Fig. S2) whereas Pavir.Bb03620 was located on a scaffold assigned to chromosome 2b (Fig S3). Furthermore, in switchgrass, Pavir.Ba00165, Pavir.Ba00166, Pavir.Ba00167 and Pavir.Ba00168 were paired with Pavir.Ba03620, Pavir.Bb03604, Pavir.Bb03602 and Pavir.Bb03603 (Fig. S3), respectively, which likely represent homeologs from the 'B' genome.

Subsets of orthologous Clade 3 peroxidases are upregulated in switchgrass and foxtail millet under aphid pressure

Of the four switchgrass peroxidases, one gene, namely Pavir.Ba00167, was expressed at significantly higher (13-fold; $P \leq 0.05$) levels in aphid-infested plants 10 DAI as compared with non-infested control plants (Fig. 3*a*). Expression levels for Pavir.Ba00165, Pavir.Ba00166 and Pavir.Ba00168 were indistinguishable between the two treatments (Fig. 3*a*). Notably, although Pavir.Ba00168 is the putative switchgrass orthologue of *SbPx-1*, it was not significantly induced in aphid-infested plants. However, Pavir.Ba00167 is orthologous to Sobic.002G416600, which was highly induced in greenbug infested sorghum signifying that these orthologs could have similar functions in defence in both sorghum and switchgrass.

Of the seven foxtail millet peroxidases, four peroxidases were significantly upregulated in greenbug-infested plants. Most notable were the three peroxidases Si030339, Si030562 and Si030595, which had average fold change values above 15 ($P \leq 0.05$) (Fig. 3*b*). The predicted orthologue of *SbPx-1* (Si030597) did not respond to aphid feeding at 10 DAI, indicating that, like the switchgrass orthologue (Pavir.Ba00168), it may occupy a different physiological niche in foxtail millet or its expression was not induced upon aphid feeding in this variety of foxtail millet (Fig. 3*b*). Notably, Si030339, which is the orthologue of aphid responsive genes Sobic.002G416600 and Pavir.Ba00168, was highly induced 10 DAI (21-fold; $P \leq 0.05$)


(a)
Fold Change
60
50
40
30
20
10
0
*
*
*
Sobic. 002G416500
Sobic. 002G416600
Sobic. 002G416700
Sobic. 002G416800
Sobic. 002G416900
Gene ID
(b)
Sobic.002G416900
Pavir.Ba00165
Si030541
Sobic.002G357000
Si030595
Sobic.002G416700 SbPx-1
Pavir.Ba00168
Si030593
Si030597
Sobic.002G416600
Pavir.Ba00167
Si030339
Sobic.002G416800
Si030562
Sobic.002G416500
Pavir.Ba00166
Si030426
Sobic.002G234200
Sobic.002G307900
Sobic.002G258300
Sobic.002G010700
Sobic.002G003700
Sobic.002G003500
Sobic.002G003600
Sobic.002G003200
Sobic.002G003400
Sobic.002G391900
Sobic.002G391200
Sobic.002G391500
Sobic.002G391400
Sobic.002G391300
Sobic.002G392000
Sobic.002G392100
Sobic.002G392300
Sobic.002G392200
Sobic.002G391800
Sobic.002G391700
Sorghum chromosome 2
Peroxidases
0.3

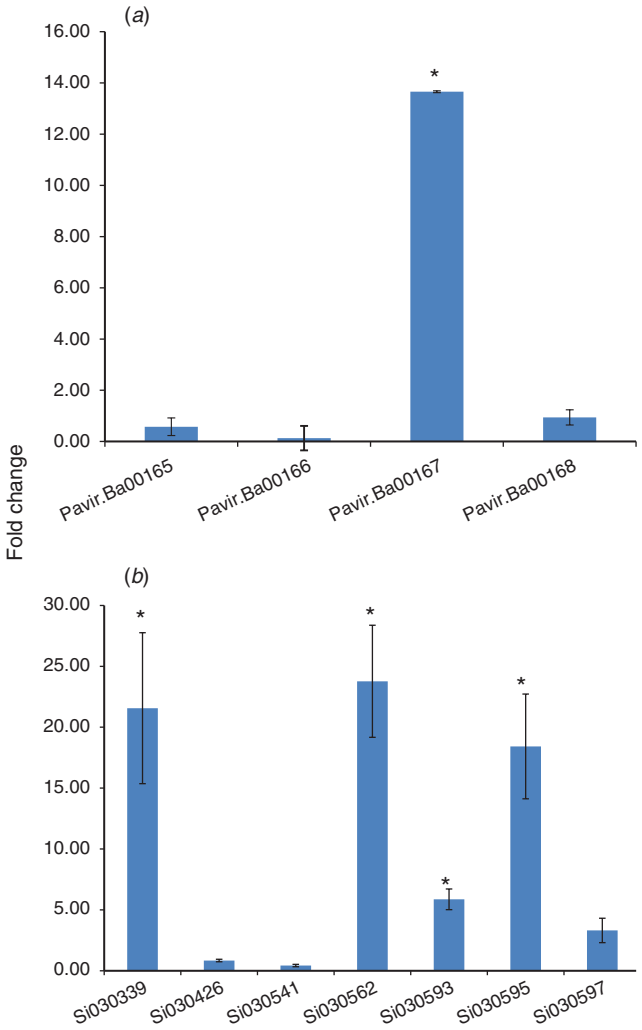


Fig. 3. Expression analysis of Clade III peroxidases in switchgrass and foxtail millet plants 10 days after aphid infestation (DAI). (*a*) RT-qPCR analysis of four peroxidase genes in switchgrass in response to greenbug feeding at 10 DAI. GAPDH was used as an endogenous control. Values are the means ± s.e. ($n = 4$). Significance of aphid infested plants relative to gene expression of the control uninfested samples is indicated: *, $P < 0.05$. (*b*) Analysis of seven peroxidase genes in foxtail millet in response to greenbug feeding at 10 DAI. Act2 was used as an endogenous control. Values are the means ± s.e. ($n = 4$). Significance of aphid infested plants relative to gene expression of the control uninfested samples at day 10 is indicated: *, $P < 0.05$.

potentially signalling that this orthologue group likely evolved to serve roles in plant defenses.

Predicted three-dimensional structures of C_4 grass Clade 3 peroxidases are indistinguishable

In a previous analysis, *SbPrx-1* was induced in sorghum upon exposure to fungal pathogens (Zamora *et al*. 2009). Eight amino acid residues, including four residues occurring in the E helix domain adjacent to the active site, were identified as being under positive selective pressure (Zamora *et al*. 2009). These residues, QxxxSLxS, were found in positions 164–171 in SbPx-1 and were hypothesised to impact the catalytic activity of the enzyme as well as the ability to bind calcium cofactors. In order to ascertain whether these residues were also present in the E-helix regions of the aphid-responsive peroxidases, the protein structures for the Clade 3 peroxidases were generated via I-Tas. SER. The structure of SbPx-1 is shown in dark blue and six of the positively selected residues are labelled (Fig. 4*a*). The backbone ribbons of all of the other models are superimposed and show very little predicted structural difference with respect to SbPx-1 (Fig. 4*a*). All of the peroxidases contained the E helix except for the aphid responsive gene Pavir.Ba00167. The model for this gene is shown in Fig. 4*b* superimposed with the model for SbPx-1 with a red box denoting the position of the E helix. In all peroxidases containing the E helix, the helix was highly variable in terms of amino acid composition, including those that were not responsive to aphid feeding (Fig. 4*c*). The QxxSLxS motif was unique to Sobic.002G416700 and was not found in any other peroxidases; moreover, this amino acid sequence in these positions did not seem to impact the structure of the E helix nor did it substantially modify the properties of the amino acid residues that comprise this motif (Fig. 4*c*). Furthermore, the amino acid sequences of the E helix region had no major impacts on the overall protein structure and did not appear to impact the catalytic residues or active site. The amino acid sequences in this region could not be conclusively linked to peroxidase responsiveness to aphid feeding and instead, correlated with phylogenetic relatedness of the peroxidases (Fig. 4*c*). For example, though Si030597 had similar amino acids in the E helix region compared with the aphid responsive genes Si03593, Si030395 and Sobic.002G416700, it did not respond to aphid feeding. Similarly, the sequences of the E-helix regions were identical in the aphid responsive gene Sobic.002G416500 and the non-responsive gene Si030426 (ANLTAAF) (Fig. 4*c*). Finally, the E helix was completely

Fig. 2. Expression analysis of clade 3 sorghum peroxidases at 10 days after aphid infestation (DAI) and phylogentic relationships to similar peroxidases in other C_4 grasses. (*a*) RT-qPCR results of sorghum Clade 3 sorghum peroxidases at 10 DAI. To determine whether any of the other sorghum peroxidases occurring in close proximity to *SbPx-1* were also induced upon aphid feeding, RT-qPCR analysis was performed on aphid infested and control plants at 10DAI. The UCE gene was used as an endogenous control. Values are the means ± s.e. ($n = 3$). Significance of aphid infested plants relative to gene expression of the control uninfested samples is indicated: *, $P < 0.05$. (*b*) Maximum likelihood (ML) analysis of syntenic peroxidases in sorghum, switchgrass, and foxtail millet. The genomes of switchgrass and foxtail millet were searched for homologues of sorghum cluster 3 peroxidases using the Phytozome gene ancestry tool. Seven peroxidases were identified in foxtail millet and four were identified in switchgrass. These sequences were subjected to ML-based phylogenetic analysis to determine their relationships with the sorghum Clade 3 peroxidases. Numbers on the nodes represent bootstrap support values ($n = 500$) and scale bar represents the number of substitutions per site. Sorghum (Sobic) sequences shaded in green are those that were induced in response to aphid feeding at 10 DAI. Sequences in brown were non-responsive.

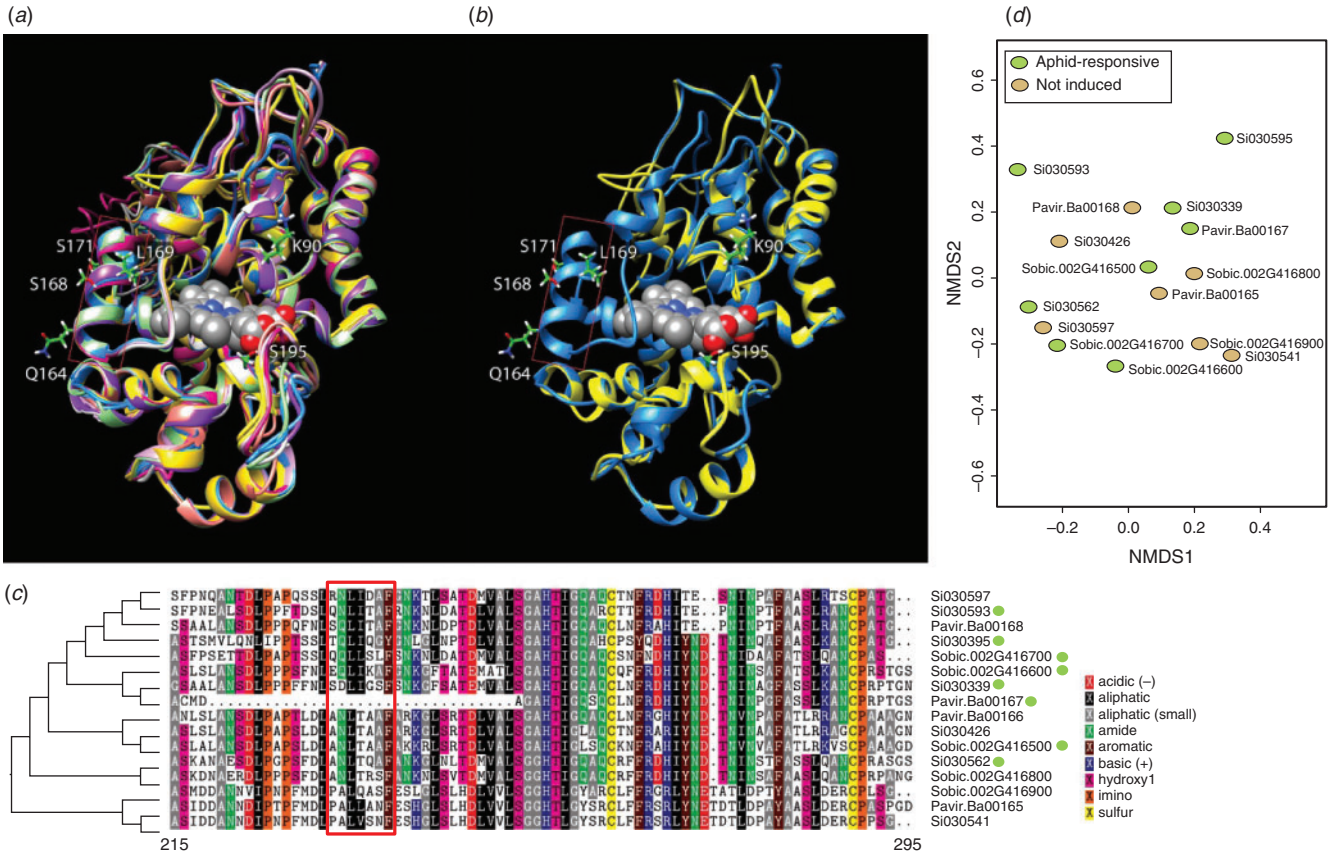


Fig. 4. Predicted structural models of aphid responsive peroxidases, amino acid sequence alignment of E helices, and NDMS analysis of *cis*-regulatory motifs in putative promoter regions of individual peroxidases. (*a*) Predicted structural models of aphid responsive peroxidases in sorghum, switchgrass, and foxtail millet. i-Tas.SER was used to generate structural predictions of peroxidases responsive to aphids using peanut peroxidase (PDB 1D:1SCH) as a template. The predicted three dimensional models for each of the 10 peroxidases were overlaid on one another using Chimera. The different coloured ribbons represent each individual peroxidase. Residues implicated as being under positive evolutionary selection in a previous study are labelled. The E helix region is highlighted with a red box. (*b*) Predicted structural model of Pavir.Ba00167. i-Tas.SER was used to generate structural predictions of Pavir.Ba00167 (yellow ribbon), which is overlaid over peanut peroxidase (PDB ID:1SCH) (blue ribbon). The E-helix region (highlighted by a red box) is missing in Pavir.Ba00167. (*c*) Multiple sequence alignment the of E-helix region of aphid responsive and non-responsive peroxidases in sorghum, switchgrass, and foxtail millet. Multiple amino acid alignment was performed using MUSCLE to determine if certain residues occurring in the E-helix region were more strongly associated with aphid-responsive peroxidases. Aphid responsive genes are indicated by green circles. (*d*) NMDS analysis of the abundances of the *cis*-regulatory elements in the promoters of the sorghum, foxtail millet and switchgrass peroxidases. To determine whether or not the composition of the *cis*-regulatory elements of the aphid-responsive genes was similar, *cis*-regulatory motifs were identified using a signal scan comparison to the PLACE database and the number of occurrences of each motif was obtained. NMDS analysis was performed using the 'vegan' library of the R statistical package. Aphid responsive genes are indicated by green circles. Brown circles indicate genes that did not respond to aphid feeding.

absent in the aphid-responsive gene Pavir.Ba00167. Future improvements to the switchgrass genome assembly will allow us to determine whether or not the missing E helix region in this gene is due to a possible misassembly in this region or if this aphid-responsive gene is truly missing the E helix domain.

Promoter analysis to identify cis*-regulatory elements in aphid responsive peroxidases*

On average, over 100 putative types of regulatory elements were identified in the promoter regions ~2000 bp upstream of the start codon in aphid-responsive peroxidases by comparison to the PLACE database. Pavir.Ba00166 occurred on a scaffold boundary and, therefore, the promoter region could not be analysed for this gene. The compositions of *cis*-regulatory elements in the promoters of aphid responsive genes were not substantially different than the uninduced peroxidase genes via NMDS analysis (Fig. 4*d*; Table S2). Furthermore, no enriched motifs (5–8 bp) were identified in the promoters of the aphid responsive genes compared with the promoter regions of genes that did not respond to aphids using the program DREME. Because the orthologue group containing Sobic.002G 416600, Pavir.Ba00167 and Si030339 was strongly induced 10 DAI in all three C_4 grass species and the promoter compositions of Si030339 and Pavir.Ba00167 were similar via NMDS analysis, the promoter regions of these three genes were also scanned for enriched motifs using DREME. Although no enriched motifs (5–8 bp) were identified in these three promoters, they all contained large numbers of CURECORECR (GTAC) and MYB2CONSENSUSAT (YAACKG) motifs, compared with

promoter elements from the majority of the other peroxidases (Table S2). Notably, the abundances of these elements were lower in promoters from peroxidases that did not respond to aphid feeding with the exception of Si030426 (14 CURCORECR motifs) and Pavir.Ba00165 (4 MYBCONSENSUSAT motifs).

Syntenic and phylogenetic comparisons supports common ancestry for Clade 3 peroxidases in C_4 grasses

To further investigate the relatedness of these peroxidases to class III peroxidases found in other grasses, a maximum likelihood based phylogenetic analysis of homologues of *SbPrx-1* found in sorghum, foxtail millet, switchgrass, *Brachypodium*, rice and maize was constructed (Fig. 5). The homologues of *SbPx-1* peroxidases were also found in genomic regions containing multiple peroxidases arranged in tandem arrays in maize (three peroxidases), rice (six peroxidases) and *Brachypodium* (five peroxidases). Notably, an orthologue of the potential progenitor peroxidase was detected in the genomes of all grass species examined (Clade 3a), suggesting it appeared before the divergence of C_3 and C_4 plants (Fig. 5). Notably, 1 : 1 : 1 : 1 : 1 orthologous sequences in the clade containing the aphid responsive genes Pavir.Ba00167 and Si030339 and Sobic.002G416600 (Clade 3f) were detected in maize, rice and *Brachypodium* indicating that this particular gene was also present before the divergence of C_3 and C_4 grass species (Fig. 5). The remainder of the peroxidases in the rice and *Brachypodium* genomes were present in either species specific clades (Clade 3d) or C_3 grass-specific clades (Clade 3e), indicating that they either appeared in the genome after the divergence of C_3 and C_4 grasses or that they have diversified so significantly from their C_3 or C_4 grass orthologs that their phylogenetic signatures have been lost (Fig. 5). Further, a clade of peroxidase specific to C_4 grasses that included a single switchgrass peroxidase (Pavir.Ba00168), three foxtail millet peroxidases (Si030593, Si030595 and Si030597), and a single sorghum peroxidase (*SbPx-1*) was also apparent. Three of the five peroxidases that were associated with this clade responded

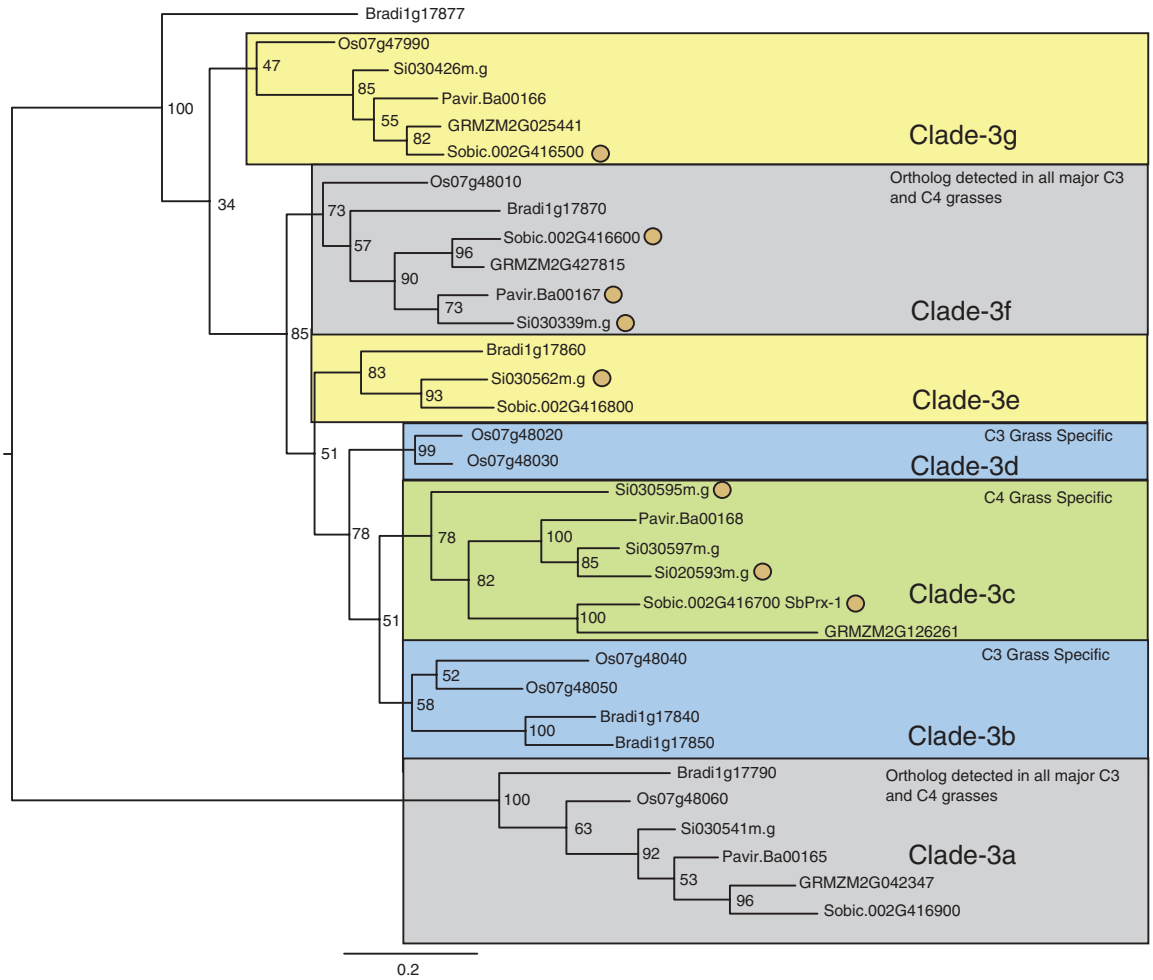


Fig. 5. Maximum likelihood analysis to identify homologues of sorghum Clade 3 peroxidases in C_3 and C_4 grasses. The genomes of rice, maize, *Brachypodium*, switchgrass and Setaria were searched for homologues of sorghum Clade 3 peroxidases using the Phytozome gene ancestry tool. Phylogenetic trees were constructed using Garli with 500 bootstrap pseudoreplications. Clades 3a and 3f have orthologs in all major C_3 and C_4 grasses whereas Clade 3c is specific to C_4 grasses and clades 3b and 3d are specific to C_3 grasses. Values at branch points denote bootstrap support and scale bar denotes number of substitutions per site.

to aphid feeding; however, Si030597 and Pavir.Ba00168, which are the nearest foxtail millet and switchgrass orthologs to *SbPx-1*, were not induced 10 DAI (Fig. 2*b*). The presence of three foxtail millet peroxidases in this clade suggests that lineage-specific expansions occurred in the foxtail millet genome. In addition, the nearest switchgrass and foxtail millet orthologs to *SbPrx-1* were not induced in response to aphid feeding indicating that these genes may serve other roles in plant physiology.

In addition, an in-depth assessment of the gene order in this region revealed that the gene order in the flanking regions are also highly conserved in C_4 and C_3 grasses, providing further evidence that these peroxidases share common evolutionary origins (Fig. S4). In both sorghum and foxtail millet, the array of peroxidases is flanked by orthologous genes and the order is largely conserved in both the 5′ and 3′ regions. In switchgrass, the 5′ and 3′ flanking regions also contain the same orthologs, but these genes are separated from the peroxidase cluster by ~30 genes (~650 kb) on the 5′ end (Fig. S4). On the 3′ end, two of the five conserved orthologs can be found two genes downstream of the peroxidase array (~36.8 kb) while the other three orthologs were found 51 genes (~532 kb) downstream of the peroxidase array (Fig. S4). Orthologs were readily identified in the 5′ flanking region in *Brachypodium* and the 3′ flanking regions in rice (Fig. S5*a*) and maize (Fig. S5*b*).

Sobic.002G416700 (SbPrx-1) is highly induced in a resistant sorghum cultivar (Cargill 607E)

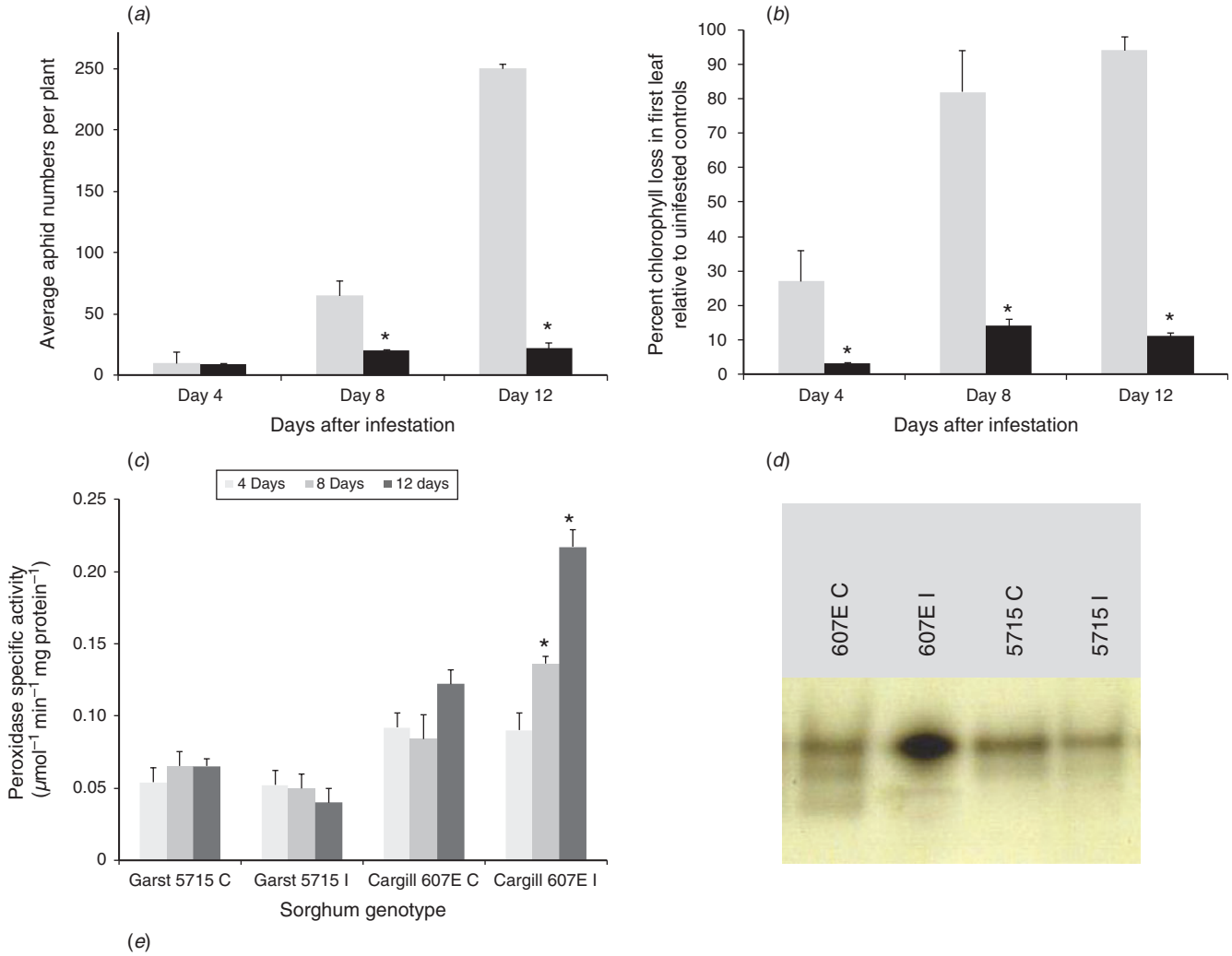
Higher numbers of aphids and higher losses of chlorophyll were observed on the greenbug susceptible Garst 5715 plants versus the resistant Cargill 607E plants (Fig. 6*a*, *b*). Among the other changes, there were differences in peroxidase activities between these sorghum hybrids as a response to aphid feeding. Greenbug-infested Cargill 607E plants consistently had higher peroxidase levels. These activities were significantly enhanced ($P \leq 0.05$) at 8 and 12 DAI in infested Cargill 607E plants as compared with activities present in uninfested control plants (Fig. 6*c*). No significant differences in peroxidase activities were observed at any harvest dates between the infested and control plants from the susceptible Garst 5715 genotype. Native gel electrophoresis of ~20 µg total protein from different plant extracts indicated that at least one isozyme was significantly upregulated upon infestation in Cargill 607E plants as compared with the susceptible Garst 5715 plants at 12 DAI (Fig. 6*d*). This isozyme also was detected in uninfested plants, albeit at lower levels (Fig. 6*d*). *De novo* peptide sequencing identified a peptide that matched exactly to a sorghum peroxidase encoded by Sobic.002G416700 (*SbPx-1*). The peptide identified by mass spectrometry is shown (shaded) on the protein sequence (Fig. 6e).

Discussion

Class III peroxidases have been implicated in many different aspects of plant defence, including the deposition of defence lignin and protection of the cell membrane against damage from ROS (Passardi *et al*. 2005; Saathoff *et al*. 2013) and the involvement of specific peroxidases in response to biotic stresses have been documented in several plant species (Dowd and Lagrimini 2006; Zamora *et al*. 2009; Suzuki *et al*. 2012). Over 100 peroxidase genes are encoded in the genomes of most grasses; however, readily distinguishing defence-related peroxidases from those involved in other metabolic pathways such as the biosynthesis of lignin or the detoxification of H_2O_2 from chloroplasts or mitochondria, is not possible based simply on nucleotide or amino acid sequences. Orthologous peroxidases occurring in syntenic regions in multiple grass genomes may contain genes that perform similar physiological roles.

The findings of this study demonstrate that synteny and orthology analyses can be used to facilitate the identification of genomic intervals containing genes involved in defence responses. For example, the genomic regions containing the Clade 3 sorghum peroxidases and their switchgrass and foxtail millet orthologs represent a genomic hot spot for peroxidases that respond to biotic stresses as several were upregulated in response to aphid feeding in these three plants. Further, the identification of a single orthologue group that includes 1 : 1 : 1 orthologous peroxidases in switchgrass (Pavir.Ba00167), foxtail millet (Si030339) and sorghum (Sobic.002G416600) that respond to aphid feeding at 10 DAI indicates that, in some cases, orthology can be an indicator of common biological function, even within multi-gene families whose evolution was largely shaped by gene amplification events. Furthermore, the rice orthologue (Os07 g48010) from this clade was induced upon exposure to the rice blast fungus *Magnaporthe oryzae* and has been linked to resistant responses to fungal pathogens in a previous study (Gupta *et al*. 2012). This observation, along with the strong induction of switchgrass, sorghum and foxtail millet orthologs after exposure to aphids, strongly suggests that this peroxidase appeared in the least common ancestor of C_3 and C_4 plants and likely evolved as a defence mechanism against biotic stresses.

Although this genomic interval contains large numbers of peroxidase genes induced in response to greenbug feeding at 10 DAI and at least one orthologue group that conclusively responds to biotic stresses in at least four grass species, orthology was not always indicative of responsiveness to aphid feeding. For example, a C_4 grass-specific clade (Clade 3c) that contained three aphid-responsive peroxidases was identified. This clade contained SbPx-1, its switchgrass orthologue (Pavir.Ba00168), its foxtail millet orthologue (Si030597), its maize orthologue (pmPOX3–2) and two other foxtail millet peroxidases (Si030595 and Si030593) that are directly adjacent to Si030597 on foxtail millet chromosome 2. Three of the six genes in this clade were induced via aphid feeding, and the maize gene was previously shown to respond to pathogen elicitors (Mika *et al*. 2010), suggesting that this orthologue group predominantly contains genes that respond to biotic stressors. However, the foxtail millet and switchgrass orthologs (Si030597 and Pavir.Ba00168) of *SbPx-1* did not respond to aphid feeding at 10 DAI even though the two other foxtail millet peroxidases assigned to this clade were induced 10 DAI. In the case of foxtail millet, the induction of other co-orthologous peroxidases (Si030593 and Si030595) in this clade indicate that lineage specific expansions of this orthologue group occurred in the genome of foxtail millet and that several of these co-orthologs have the ability to respond to biotic stresses in foxtail millet. In the case of switchgrass, it is possible that the ‘B’ genome homeolog of Pavir.Ba00168 (Pavir. Bb03602) could respond to aphid feeding. This observation



(*e*)

1 MAASASCLA FLVAAAVASV ASAQLSSTFY DTSCPNALST IKSGVDAAVM QEARTGASLLR 61
MHFHDCFVH GCDGSVLLND TSGEQSSPPN KGSLRRFDVI DSIKAQVEAV CPGVVSCADIL 121
AVAARDSVV ALGGPSWTVL LGRRDSTASF PSETTDLPAP TSSLQQLLSL FSNKNLDATDM 181
VALSGAHTI GQAQCSNFND HIYNDTNIDA AFATSLQANC PASGSTSLAP LDTMTPTTFDN 241
DYYTNLMSQ KGLLHSDQEL FNNGSTDSTV SNFASSASAF TSAFTAAMVK MGNLSPLTGTD
301 GEIRLACGI VNSS

Fig. 6. Greenbug aphid numbers, leaf chlorophyll content, peroxidase activities, and identification of SbPx-1 by gel electrophoresis and mass spectrometry. (*a*) Aphid numbers and (*b*) leaf chlorophyll content in greenbug susceptible Garst 5715 (grey bars) and greenbug resistant Cargill 607E (black bars) sorghum plants at 4, 8 and 12 days after aphid infestation (DAI). In (*b*) chlorophyll loss is expressed as a function of leaf chlorophyll content in uninfested control plants. (*c*) Peroxidase specific activities in control (C) and infested (I) plants. Significant differences between infested and control plants are indicated: *. $P \leq 0.05$. (*d*) Native gel stained for peroxidase activity at 12 DAI. (*e*) Amino acid sequence of SbPx-1 (Sobic.002G416700). The peptide identified by mass spectrometry is shaded. Significant differences between infested and control plants are indicated: *, $P \leq 0.05$.

indicates that the identification of syntenic regions in multiple grass species can facilitate the identification of genomic intervals with the ability to respond to aphid feeding and other biotic stresses, but that additional functional characterisation may be required to identify specific genes in these intervals that respond to biotic stress.

Further, genes occurring within this genomic interval are linked with resistant responses in sorghum. For example, at least one specific peroxidase isozyme (SbPrx-1) accumulated to high levels in tissues collected from resistant Cargill 607E plants in comparison to susceptible Garst 5715 lines at 12 DAI, coincident with increased peroxidase activities observed in plant extracts. However, the actual contribution of SbPrx-1 to these increases in activities remains to be established. Elevated peroxidase levels in resistant plants could play a role in the detoxification of excess H_2O_2 produced in response to aphid injury (Mai *et al*. 2013) and/or contribute to cellular defenses for strengthening cell walls through lignification (Passardi *et al*.

2004*b*; Saathoff *et al*. 2013). Increased lignification in cell walls may help prevent the loss of nutrients from within the cell or protect the cell from digestive enzymes produced by insect herbivores (Furch *et al*. 2015), and the prolonged ability to detoxify H_2O_2 may enable resistant plants to sustain elevated ROS levels for extended periods of time without incurring major damage to cellular components. In many cases, responses to pathogens and insect herbivory share overlapping mechanisms, such as an upregulation of genes responsive to methyl jasmonate, ethylene and ROS (Gutsche *et al*. 2009*a*; Lazebnik *et al*. 2014). Supporting this hypothesis, SbPrx-1 was associated with resistance to *Colletotrichum sublineolum*, the causal agent of anthracnose (Pratt *et al*. 2005) and was determined to be under positive selective pressure (Zamora *et al*. 2009) and pmPOX3–2, the maize orthologue of *SbPrx-1*, is induced by methyl jasmonate, salicylic acid and pathogen elicitors (Mika *et al*. 2010), suggesting it could also serve roles in plant defenses. Although SbPrx-1 is associated with resistant responses to aphids and fungal pathogens in sorghum, more research is needed to determine whether other peroxidases in this interval or elsewhere in the genome could also be linked to resistant responses to biotic stresses in sorghum and other grasses. In addition, the expression of some of these peroxidases could be under circadian control, and low expression values for specific genes under aphid infestation could be reflective of the harvest time. Further study is needed to determine if any of these peroxidases are under circadian control and if they also could be linked to responses to biotic stresses.

Although several genes occurring in this genomic interval respond similarly to aphid pressure in sorghum, switchgrass and foxtail millet, their mechanism of induction and the identification of cellular cues that trigger their responses to biotic stresses are not currently known. NMDS analysis failed to conclusively separate the promoters of the aphid responsive genes from those that did not respond to herbivory, indicating that the compositions of the promoter elements are similar. Further, no novel enriched motifs were detected in the aphid responsive genes using DREME and no motifs that were exclusively associated with the aphid-induced peroxidases could be identified; however, MYB2CONSENSUSAT and CURECORECR motifs were highly abundant in the promoters of the orthologue group of aphid responsive genes Sobic.002G416600, Pavir.Ba00167 and Si030339 and may be linked to their activation under periods of biotic stress. Several myb transcription factors induce gene expression under periods of biotic stress in grasses and activate networks of defence-related pathways (Ibraheem *et al*. 2015; Zhang *et al*. 2015) and CURECORECR motifs are responsive to changes in oxygen levels and may be triggered by changes in ROS levels that often accompany insect feeding. Other regulatory elements that might induce the expression of all or a subset of the peroxidases in this interval include long range promoters or enhancers (Sanyal *et al*. 2012). Another possibility is that transcription and regulatory factors responsible for activating these genes bind to highly degenerate motifs, which would not be detected in this analysis (Bailey *et al*. 2006). The improvement of the switchgrass genome assembly and the future release of other C_3 and C_4 grass genomes will enable us to assess whether or not these genes are conserved and respond similarly to biotic pressure in all grasses.

Conclusions

This research has identified a syntenic region present in the genomes of all sequenced C_3 and C_4 grasses that contains several orthologous peroxidases, several of which are upregulated in sorghum, switchgrass and foxtail millet plants in response to greenbug feeding. Additionally, some of the rice and maize peroxidases occurring in this interval have been previously shown to respond to pathogens, suggesting that this region is a hotspot for peroxidases that respond to biotic stresses. Further, at least one C_4 grass-specific peroxidase was highly induced in the resistant sorghum cultivar Cargill 607E indicating that this region could be linked to resistant responses in sorghum and other grasses. Despite the apparent lack of common *cis*-elements in the peroxidase genes, the universal responsiveness of these orthologous genes to aphid feeding across several grass species indicates that common mechanisms of defence response may exist in C_4 grasses and that several of these defence responses may even be common to both C_3 and C_4 grasses. These genes can serve as markers for improving aphid tolerance in the three C_4 grasses examined and they can be explored as resistance markers in other C_3 and C_4 grasses species.

Acknowledgements

We gratefully acknowledge Dana Novak, Andrea Gutsche, and Nathan Palmer for technical assistance. We also thank Dr Ishmail Dweikat (University of Nebraska-Lincoln, Lincoln, NE) for providing the foxtail millet seeds and Dr Robert A Graybosch (USDA-ARS) for his comments to improve the manuscript. This research was supported in part by the University of Nebraska Agricultural Experiment Station Project NEB-41–041, Nebraska Sorghum Board, the Office of Science (BER), USA Department of Energy Grant Number DE-AI02–09ER64829, USDA-NIFA Grant Numbers 2011–67009–30096, 2011–67009–30026, and by the USDA-ARS CRIS project 5440–21000–030–00D. The USA Department of Agriculture, Agricultural Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organisations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organisations not mentioned.

References

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105. doi:10.1093/bioinformatics/bti263

Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology* **55**, 539–552. doi:10.1080/10635150600755453

Bailey TL (2011) DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659. doi:10.1093/bioinformatics/btr261

Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research* **34**, W369–W373. doi:10.1093/nar/gkl198

Barbehenn R, Dukatz C, Holt C, Reese A, Martiskainen O, Salminen JP, Yip L, Tran L, Constabel CP (2010) Feeding on poplar leaves by caterpillars potentiates foliar peroxidase action in their guts and increases plant resistance. *Oecologia* **164**, 993–1004. doi:10.1007/s00442-010-1733-y

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye CY, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM (2012) Reference genome sequence of the model plant Setaria. *Nature Biotechnology* **30**, 555–561. doi:10.1038/nbt.2196

Berman H, Henrick K, Nakamura H (2003) Announcing the worldwide Protein Data Bank. *Nature Structural Biology* **10**, 980. doi:10.1038/nsb1203-980

Carmona D, Fornoni J (2013) Herbivores can select for mixed defensive strategies in plants. *New Phytologist* **197**, 576–585. doi:10.1111/nph.12023

Chaman ME, Corcuera LJ, Zuniga GE, Cardemil L, Argandona VH (2001) Induction of soluble and cell wall peroxidases by aphid infestation in barley. *Journal of Agricultural and Food Chemistry* **49**, 2249–2253. doi:10.1021/jf0011440

Distelfeld A, Pearce SP, Avni R, Scherer B, Uauy C, Piston F, Slade A, Zhao R, Dubcovsky J (2012) Divergent functions of orthologous NAC transcription factors in wheat and rice. *Plant Molecular Biology* **78**, 515–524. doi:10.1007/s11103-012-9881-6

Dixon P (2003) VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14**, 927–930. doi:10.1111/j.1654-1103.2003.tb02228.x

Dogramaci M, Mayo ZB, Wright R, Reese J (2007) Categories of resistance, antibiosis and tolerance, to biotype I greenbug (*Schizaphis graminum* (Rondani) Homoptera: Aphididae) in four sorghum (*Sorghum bicolor* (L.) Moench. Poales: Gramineae) hybrids. *Journal of the Kansas Entomological Society* **80**, 183–191. doi:10.2317/0022-8567(2007)80[183:CORAAT]2.0.CO;2

Dowd PF, Lagrimini LM (2006) Examination of the biological effects of high anionic peroxidase production in tobacco plants grown under field conditions. I. Insect pest damage. *Transgenic Research* **15**, 197–204. doi:10.1007/s11248-005-3913-9

Dowd PF, Holmes RA, Pinkerton TS, Johnson ET, Lagrimini LM, Boston RS (2006) Relative activity of a tobacco hybrid expressing high levels of a tobacco anionic peroxidase and maize ribosome-inactivating protein against *Helicoverpa zea* and *Lasioderma serricorne*. *Journal of Agricultural and Food Chemistry* **54**, 2629–2634. doi:10.1021/jf058180p

Franzen LD, Gutsche AR, Heng-Moss TM, Higley LG, Sarath G, Burd JD (2007) Physiological and biochemical responses of resistant and susceptible wheat to injury by Russian wheat aphid. *Journal of Economic Entomology* **100**, 1692–1703. doi:10.1093/jee/100.5.1692

Furch ACU, van Bel AJE, Will T (2015) Aphid salivary proteases are capable of degrading sieve-tube proteins. *Journal of Experimental Botany* **66**, 533–539. doi:10.1093/jxb/eru487

Goodstein DM, Shu SQ, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**, D1178–D1186. doi:10.1093/nar/gkr944

Gulsen O, Eickhoff T, Heng-Moss T, Shearman R, Baxendale F, Sarath G, Lee D (2010) Characterization of peroxidase changes in resistant and susceptible warm-season turfgrasses challenged by *Blissus occiduus*. *Arthropod-Plant Interactions* **4**, 45–55. doi:10.1007/s11829-010-9086-3

Gupta SK, Rai AK, Kanwar SS, Chand D, Singh NK, Sharma TR (2012) The single functional blast resistance gene Pi54 activates a complex defence mechanism in rice. *Journal of Experimental Botany* **63**, 757–772. doi:10.1093/jxb/err297

Gutsche A, Heng-Moss T, Sarath G, Twigg P, Xia Y, Lu G, Mornhinweg D (2009*a*) Gene expression profiling of tolerant barley in response to *Diuraphis noxia* (Hemiptera: Aphididae) feeding. *Bulletin of Entomological Research* **99**, 163–173. doi:10.1017/S0007485308006184

Gutsche AR, Heng-Moss TM, Higley LG, Sarath G, Mornhinweg DW (2009*b*) Physiological responses of resistant and susceptible barley, *Hordeum vulgare* to the Russian wheat aphid, *Diurpahis noxia* (Mordvilko). *Arthropod-Plant Interactions* **3**, 233–240. doi:10.1007/s11829-009-9067-6

Heng-Moss T, Sarath G, Baxendale F, Novak D, Bose S, Ni X, Quisenberry S (2004) Characterization of oxidative enzyme changes in buffalograsses challenged by *Blissus occiduus*. *Journal of Economic Entomology* **97**, 1086–1095. doi:10.1603/0022-0493(2004)097[1086:COOECI]2.0.CO;2

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**, 297–300. doi:10.1093/nar/27.1.297

Ibraheem F, Gaffoor I, Tan QX, Shyu CR, Chopra S (2015) A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules* **20**, 2388–2404. doi:10.3390/molecules20022388

Kayser JP, Kim JG, Cerny RL, Vallet JL (2006) Global characterization of porcine intrauterine proteins during early pregnancy. *Reproduction* **131**, 379–388. doi:10.1530/rep.1.00882

Kerchev PI, Fenton B, Foyer CH, Hancock RD (2012) Plant responses to insect herbivory: interactions between photosynthesis, reactive oxygen species and hormonal signalling pathways. *Plant, Cell & Environment* **35**, 441–453. doi:10.1111/j.1365-3040.2011.02399.x

Koch KG, Fithian R, Heng-Moss TM, Bradshaw JD, Sarath G, Spilker C (2014) Evaluation of tetraploid switchgrass (Poales: Poaceae) populations for host suitability and differential resistance to four cereal aphids. *Journal of Economic Entomology* **107**, 424–431. doi:10.1603/EC13315

Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**, 299–306. doi:10.1093/bib/bbn017

Lazebnik J, Frago E, Dicke M, van Loon JJA (2014) Phytohormone mediation of interactions between herbivores and plant pathogens. *Journal of Chemical Ecology* **40**, 730–741. doi:10.1007/s10886-014-0480-7

Lei JX, Finlayson SA, Salzman RA, Shan LB, Zhu-Salzman K (2014) BOTRYTIS-INDUCED KINASE1 modulates *Arabidopsis* resistance to green peach aphids via PHYTOALEXIN DEFICIENT4. *Plant Physiology* **165**, 1657–1670. doi:10.1104/pp.114.242206

Louis J, Shah J (2013) *Arabidopsis thaliana-Myzus persicae* interaction: shaping the understanding of plant defense against phloem-feeding aphids. *Frontiers in Plant Science* **4**, 213. doi:10.3389/fpls.2013.00213

Mai VC, Bednarski W, Borowiak-Sobkowiak B, Wilkaniec B, Samardakiewicz S, Morkunas I (2013) Oxidative stress in pea seedling leaves in response to *Acyrthosiphon pisum* infestation. *Phytochemistry* **93**, 49–62. doi:10.1016/j.phytochem.2013.02.011

Marchi-Werle L, Heng-Moss TM, Hunt TE, Baldin ELL, Baird LM (2014) Characterization of peroxidase changes in tolerant and susceptible soybeans challenged by soybean aphid (Hemiptera: Aphididae). *Journal of Economic Entomology* **107**, 1985–1991. doi:10.1603/EC14220

Massa AN, Wanjugi H, Deal KR, O'Brien K, You FM, Maiti R, Chan AP, Gu YQ, Luo MC, Anderson OD, Rabinowicz PD, Dvorak J, Devos KM (2011) Gene space dynamics during the evolution of *Aegilops tauschii*, *Brachypodium distachyon*, *Oryza sativa*, and *Sorghum bicolor* genomes. *Molecular Biology and Evolution* **28**, 2537–2547. doi:10.1093/molbev/msr080

Michels GJ, Burd JD (2007) IPM case studies: sorghum. In 'Aphids as crop pests'. (Eds HF Emden, R van Harrington) pp. 627–637. (CABI International: Walllingford, UK)

Mika A, Boenisch MJ, Hopff D, Luthje S (2010) Membrane-bound guaiacol peroxidases from maize (*Zea mays* L.) roots are regulated by methyl jasmonate, salicylic acid, and pathogen elicitors. *Journal of Experimental Botany* **61**, 831–841. doi:10.1093/jxb/erp353

Moloi MJ, van der Westhuizen AJ (2008) Antioxidative enzymes and the Russian wheat aphid (*Diuraphis noxia*) resistance response in wheat (*Triticum aestivum*). *Plant Biology* **10**, 403–407. doi:10.1111/j.1438-8677.2008.00042.x

Ni X, Quisenberry SS, Heng-Moss T, Markwell J, Sarath G, Klucas R, Baxendale F (2001) Oxidative responses of resistant and susceptible cereal leaves to symptomatic and nonsymptomatic cereal aphid (Hemiptera: Aphididae) feeding. *Journal of Economic Entomology* **94**, 743–751. doi:10.1603/0022-0493-94.3.743

Ostlund G, Schmitt T, Forslund K, Kostler T, Messina DN, Roopra S, Frings O, Sonnhammer ELL (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Research* **38**, D196–D203. doi:10.1093/nar/gkp931

Padmaja PG, Shwetha BL, Swetha G, Patil JV (2014) Oxidative enzyme changes in sorghum infested by shoot fly. *Journal of Insect Science* **14**, 193. doi:10.1093/jisesa/ieu055

Palmer NA, Saathoff AJ, Waters BM, Donze T, Heng-Moss TM, Twigg P, Tobias CM, Sarath G (2014) Global changes in mineral transporters in tetraploid switchgrasses (*Panicum virgatum* L.). *Frontiers in Plant Science* **4**, 549. doi:10.3389/fpls.2013.00549

Passardi F, Longet D, Penel C, Dunand C (2004*a*) The class III peroxidase multigenic family in rice and its evolution in land plants. *Phytochemistry* **65**, 1879–1893. doi:10.1016/j.phytochem.2004.06.023

Passardi F, Penel C, Dunand C (2004*b*) Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends in Plant Science* **9**, 534–540. doi:10.1016/j.tplants.2004.09.002

Passardi F, Cosio C, Penel C, Dunand C (2005) Peroxidases have more functions than a Swiss army knife. *Plant Cell Reports* **24**, 255–265. doi:10.1007/s00299-005-0972-6

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612. doi:10.1002/jcc.20084

Porra RJ, Thompson WA, Kriedemann PE (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls *a* and *b* extracted with 4 different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta* **975**, 384–394. doi:10.1016/S0005-2728(89)80347-0

Pratt LH, Liang C, Shah M, Sun F, Wang H, Reid SP, Gingle AR, Paterson AH, Wing R, Dean R, Klein R, Nguyen HT, Ma HM, Zhao X, Morishige DT, Mullet JE, Cordonnier-Pratt MM (2005) Sorghum expressed sequence tags identify signature genes for drought, pathogenesis, and skotomorphogenesis from a milestone set of 16,801 unique transcripts. *Plant Physiology* **139**, 869–884. doi:10.1104/pp.105.066134

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725–738. doi:10.1038/nprot.2010.5

Saathoff AJ, Donze T, Palmer NA, Bradshaw J, Heng-Moss T, Twigg P, Tobias CM, Lagrimini M, Sarath G (2013) Towards uncovering the roles of switchgrass peroxidases in plant processes. *Frontiers in Plant Science* **4**, 202. doi:10.3389/fpls.2013.00202

Sanyal A, Lajoie BR, Jain G, Dekker J (2012) The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113. doi:10.1038/nature11279

Schnable JC (2015) Genome evolution in maize: from genomes back to genes. *Annual Review of Plant Biology* **66**, 329–343. doi:10.1146/annurev-arplant-043014-115604

Schnable JC, Freeling M, Lyons E (2012) Genome-wide analysis of syntenic gene deletion in the grasses. *Genome Biology and Evolution* **4**, 265–277. doi:10.1093/gbe/evs009

Scully ED, Gries T, Sarath G, Palmer NA, Baird L, Serapiglia MJ, Dien BS, Boateng AA, Ge Z, Funnell-Harris DL, Twigg P, Clemente TE, Sattler SE (2016) Overexpression of *SbMyb60* impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. *The Plant Journal* **85**, 378–395. doi:10.1111/tpj.13112

Suzuki H, Dowd PF, Johnson ET, Hum-Musser SM, Musser RO (2012) Effects of elevated peroxidase levels and corn earworm feeding on gene expression in tomato. *Journal of Chemical Ecology* **38**, 1247–1263. doi:10.1007/s10886-012-0205-8

Thaler JS, Humphrey PT, Whiteman NK (2012) Evolution of jasmonate and salicylate signal crosstalk. *Trends in Plant Science* **17**, 260–270. doi:10.1016/j.tplants.2012.02.010

Wu J, Wang L, Wunsche H, Baldwin IT (2013) Narboh D, a respiratory burst oxidase homolog in *Nicotiana attenuata*, is required for late defense responses after herbivore attack. *Journal of Integrative Plant Biology* **55**, 187–198. doi:10.1111/j.1744-7909.2012.01182.x

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**, 1586–1591. doi:10.1093/molbev/msm088

Zamora A, Sun Q, Hamblin MT, Aquadro CF, Kresovich S (2009) Positively selected disease response orthologous gene sets in the cereals identified using *Sorghum bicolor* L. Moench expression profiles and comparative genomics. *Molecular Biology and Evolution* **26**, 2015–2030. doi:10.1093/molbev/msp114

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40. doi:10.1186/1471-2105-9-40

Zhang ZJ, Chen JM, Su YY, Liu HM, Chen Y, Luo PG, Du XG, Wang D, Zhang HY (2015) *TaLHY*, a 1R-MYB transcription factor, plays an important role in disease resistance against stripe rust fungus and ear heading in wheat. *PLoS One* **10**, e0127723. doi:10.1371/journal.pone.0127723

Zwicki DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, School of Biological Sciences, University of Texas-Austin.