New Phytologist

Evidence of an evolutionary hourglass pattern in herbivory-induced transcriptomic responses

Matthew Durrant[1,2], Justin Boyer[1,2], Wenwu Zhou[1], Ian T. Baldwin[1] and Shuqing Xu[1]

[1]Department of Molecular Ecology, Max Planck Institute for Chemical Ecology, Hans-Knöll-Str. 8, 07745 Jena, Germany; [2]Department of Plant and Wildlife Sciences, Brigham Young University, 4105A, LSB, Provo, UT 84602, USA

Author for correspondence:
Shuqing Xu
Tel: +49 0 3641 571122
Email: sxu@ice.mpg.de

Received: *8 December 2016*
Accepted: *28 April 2017*

New Phytologist (2017) **215:** 1264–1273
doi: 10.1111/nph.14644

Key words: defense signaling, evolutionary hourglass, herbivory-induced defense, *Nicotiana attenuata*, phylotranscriptomic analysis, primary metabolism, specialized metabolism.

Summary

• Herbivory-induced defenses are specific and activated in plants when elicitors, frequently found in the herbivores' oral secretions, are introduced into wounds during attack. While complex signaling cascades are known to be involved, it remains largely unclear how natural selection has shaped the evolution of these induced defenses.

• We analyzed herbivory-induced transcriptomic responses in wild tobacco, *Nicotiana attenuata*, using a phylotranscriptomic approach that measures the origin and sequence divergence of herbivory-induced genes.

• Highly conserved and evolutionarily ancient genes of primary metabolism were activated at intermediate time points (2–6 h) after elicitation, while less constrained and young genes associated with defense signaling and biosynthesis of specialized metabolites were activated at early (before 2 h) and late (after 6 h) stages of the induced response, respectively – a pattern resembling the evolutionary hourglass pattern observed during embryogenesis in animals and the developmental process in plants and fungi.

• The hourglass patterns found in herbivory-induced defense responses and developmental process are both likely to be a result of signaling modularization and differential evolutionary constraints on the modules involved in the signaling cascade.

Introduction

Herbivory-induced defenses are widespread in plants and play an important role in maintaining plant fitness when they are under attack (Karban & Baldwin, 1997). The molecular mechanisms and ecological functions of herbivory-induced signaling cascades and defense responses have been examined in several plant systems (Karban & Baldwin, 1997; Bodenhausen & Reymond, 2007; Howe & Jander, 2008; Wu & Baldwin, 2010; Erb *et al.*, 2015; Li *et al.*, 2015b). After attack, chemical cues (herbivory-associated elicitors, HAEs) in insect oral secretions (OSs) elicit a series of signaling cascades and induced defenses in plants, which directly or indirectly deter feeding herbivores and protect plants from further damage (Turlings *et al.*, 1990; Mattiacci *et al.*, 1995; Karban & Baldwin, 1997; Agrawal, 1998; Kessler *et al.*, 2004; Unsicker *et al.*, 2009). For example, in a wild tobacco species, *Nicotiana attenuata*, the fatty acid amino acid conjugates (FACs) found in the OSs of *Manduca sexta* – an insect that largely feeds on solanaceous plants – elicit rapid phytohormonal changes in leaves, including the accumulation of jasmonic acid (JA) and its derivatives when OSs are introduced into wounds as larvae feed on leaves (Wu & Baldwin, 2010; Bonaventure *et al.*, 2011). The amplification of the wound-induced JA burst by OSs activates the biosynthesis of several potent herbivore toxins, such as phenolamides and 17-hydroxygeranyllinalool diterpene glycosides (HGL-DTGs) (Gulati *et al.*, 2013; Gaquerel *et al.*, 2014), which function as antiherbivore defenses in tests conducted in both the laboratory and the native environment of *N. attenuata* (Heiling *et al.*, 2010; Kaur *et al.*, 2010; Gaquerel *et al.*, 2014).

Despite the advantages provided by induced defenses, they can also have negative effects on plant fitness as a result of their physiological and ecological costs (Baldwin, 1998; van Dam & Baldwin, 1998; Baldwin & Hamilton, 2000; Redman *et al.*, 2001; Agrawal *et al.*, 2002; Heil & Baldwin, 2002; Strauss *et al.*, 2002; Glawe *et al.*, 2003; Wang *et al.*, 2015). Based on studies that measured fitness consequences of induced defenses against herbivores, a meta-analysis showed that the fitness cost associated with induced defenses is substantial (52%, $n = 23$) (Strauss *et al.*, 2002). Therefore, the defenses in different plant populations and species that grow in heterogeneous environments are subject to divergent selection pressures, which are driven by the effectiveness of induced defenses against the local herbivore communities and the costs associated with the activation of the induced defenses. These divergent selection pressures are likely to account for the observed variability in induced defenses found among closely related species and populations of the same species (Li *et al.*, 2015a; Xu *et al.*, 2015; Zhou *et al.*, 2016).

Herbivory is known to elicit different transcriptomic and metabolomic responses at different times after elicitation in plants (Gulati *et al.*, 2013; Tzin *et al.*, 2015; Liu *et al.*, 2016; Qi *et al.*, 2016), suggesting that the induced defense signaling

cascade consists of distinct transcriptomic modules. How these distinct gene modules respond to selection remains unclear. Like herbivory-induced defense signaling, developmental processes also consist of complex signaling cascades which have been studied by phylotranscriptomic analysis (Domazet-Loso & Tautz, 2010; Kalinka *et al.*, 2010; Quint *et al.*, 2012; Cheng *et al.*, 2015; Drost *et al.*, 2015). This method incorporates the age of genes, which is inferred from when a gene first appears in the evolutionary lineage of the species of interest (Domazet-Loso & Tautz, 2010; Kalinka *et al.*, 2010), and the rate of sequence divergence observed within the genes, which is inferred from the protein coding sequences. Previous studies have used both the transcriptome age index (TAI), which is calculated based on the average 'age' of all expressed genes, and the transcriptome divergence index (TDI), which is calculated from the average sequence divergence of all expressed genes. High TAI values indicate that a large proportion of expressed genes are young, while high TDI values indicate that a large proportion of expressed genes are less constrained in their sequence divergence. TAI and TDI are complementary in that they represent ancient and recent selection pressures, respectively (Quint *et al.*, 2012).

In this study, we adapted the concept of phylotranscriptomic analysis to investigate the sequence of changes in gene expression that occur after herbivore elicitation. We developed two indices, the induced TAI (iTAI) and induced TDI (iTDI), which are closely analogous to the TAI and TDI, respectively. This approach allowed us to specifically analyze the average gene age and sequence divergence of the induced genes relative to a baseline untreated control throughout a time series analysis of an induced response, and thereby capture an evolutionary analysis of the gene sets recruited during the activation of herbivory-induced defenses.

Materials and Methods

Plant material and treatment, microarray data and analysis

Plant materials and treatment were the same as described previously (Kim *et al.*, 2011). In brief, seeds of wild-type (WT) *Nicotiana attenuata* Torr. Ex S. Watson plants (30th inbred generation) were sterilized and germinated on Gamborg's B5 medium as previously described (Krügel *et al.*, 2002). After 10 d, seedlings were transferred to Teku pots (Pöppelmann GmbH & Co. KG, Lohne, Germany) in the glasshouse. Plants were then transferred to 1 l pots with soil after 10 d and were grown under glasshouse conditions at 26–28°C with 16 h supplemental light from Master Sun-T PIA Agro 400 or Master Sun-T PIA Plus 600 W Na lights (Philips, Turnhout, Belgium). To simulate herbivore attack, one leaf of each rosette stage plant was wounded with a pattern wheel and 20 µl of 1 : 5 diluted *Manduca sexta* OS was added to the puncture wounds. For the control group, leaves from untreated plants were harvested. Samples were collected at 0.5, 1, 2, 5, 6, 11 and 13 h after elicitation, their midveins rapidly excised, flash-frozen in liquid nitrogen and stored at −80°C until analysis. The elicitations were performed at 13:00 h. The time points selected for sample collection were based on our previous analysis of the pilot dataset (https://doi.org/10.1101/034603). Three biological replicates were used at each time point.

Similar to a previous study (Kim *et al.*, 2011), total RNA was isolated using the TRIZOL reagent and labeled cRNA with the Quick Amp labeling kit (Agilent, Waldbronn, Germany). Each sample was hybridized on Agilent single color technology arrays (4 × 44K 60-mer oligonucleotide microarray). Quantile normalization and the $\log_2$ transformation were performed on all raw microarray data using the R package 'Agi4x44PreProcess' (http://goo.gl/TJnA6Q). Fold-change and differentially expressed genes were analyzed using the R package 'limma'.

Phylostratigraphic map

To construct the phylostratigraphic map (Supporting Information Fig. S1), we used BLASTP from the BLAST (v.2.2.25+) suite to search the curated NCBI taxonomy database (Domazet-Loso *et al.*, 2007; Domazet-Loso & Tautz, 2010) to assign *N. attenuata* genes to 13 phylostrata. This method is similar to those used in previous studies (Domazet-Loso *et al.*, 2007; Domazet-Loso & Tautz, 2010), with some modifications. In brief, all protein-coding sequences of *N. attenuata* were compared with the nonredundant (nr) NCBI protein database (downloaded on 29 April 2014) by searching BLASTP with an *E*-value cutoff of 10^{-3}. The BLASTP results were further filtered to exclude synthetic sequences, viruses, and sequences that do not descend from the 'cellular organisms' phylostratum. All genes were assigned to the phylogenetically most ancient phylostratum (PS) containing at least one species with at least one blast hit using a custom Python script. Because of the scarcity of protein sequences from the *Nicotiana* genus in the nr database, all *N. attenuata* genes without a match were further searched against a locally stored *N. obtusifolia* genome. A gene was allocated to PS 12 (*Nicotiana* specific) if a hit to *N. obtusifolia* was detected, or to PS 13 (*N. attenuata* specific) if no hit was detected. This method assumes that genes with shared domains belong to the same gene family, and therefore subsequent duplications of founder genes are generally assigned to the same PS as the founder gene, regardless of the time period in which the duplication event occurred (Domazet-Loso *et al.*, 2007).

Validation of phylostratigraphic map through search algorithm comparison

Although this method has been used previously (Domazet-Loso *et al.*, 2007; Domazet-Loso & Tautz, 2010), a recent study by Moyers and Zhang has demonstrated that using the BLASTP algorithm to find homologs can underestimate a gene's phylostratigraphic age and result in a biased phylostratigraphic map (Moyers & Zhang, 2015). To test the robustness of the hourglass pattern, we used an additional homolog iterative search algorithm, PSI-BLAST, to search distant homologs. PSI-BLAST (from BLAST v.2.2.25+) was run with a cutoff value of 10^{-3} for at least four iterations (Jones & Swindells, 2002). While the phylostratigraphic map of *N. attenuata* genes based on BLASTP resembles the distribution reported in *Arabidopsis thaliana* by Quint *et al.*

(2012), the *N. attenuata* phylostratigraphic map based on PSI-BLAST resulted in a larger number of genes assigned to earlier PS groups, as predicted by Moyers & Zhang (2015). For example, the number of genes assigned to the PS1 group increased from 4326 (BLASTP) to 6309 (PSI-BLAST). Despite these shifts in age distributions among the searching algorithms, the corresponding iTAI patterns all resulted in the same hourglass pattern, with younger genes induced at greater levels at early and late time points, while the average age of genes recruited at the intermediate time points is much older (Fig. S2).

K_a/K_s ratios

The K_a/K_s ratio was calculated using the codeml method from the PAML package (v.4.7) using ortholog genes between *N. attenuata* and *N. obtusifolia* (diverged *c.* 7 million yr ago (Ma)). The K_a/K_s ratio is calculated from aligned protein coding sequences and measures the rate of nonsynonymous substitutions per nonsynonymous nucleotide site (K_a) relative to the rate of synonymous substitutions per synonymous nucleotide site (Ks). From the K_a/K_s ratio one can infer the evolutionary constraint that a gene has been under. Genes with a K_s value < 0.05 and > 1 were removed from the downstream analysis, because manual inspection of the sequence alignments of these genes revealed that either these genes were too short and highly similar between two species, thereby violating the requirements for accurate K_a/K_s ratio calculations, or the sequence alignments were obviously incorrect. Similar to a previous study (Quint *et al.*, 2012; Drost *et al.*, 2015), our analysis showed that sequence divergence and gene age are only weakly correlated and thus can provide complementary evidence for estimations of evolutionary distance (Fig. S3; Table S1).

Transcriptome induction indices

The transcriptome indices for *N. attenuata* transcriptomic responses were initially calculated based on a previously published, publicly available microarray dataset of locally treated leaves, systemic leaves, and roots at six time points within 21 h of simulated herbivore attack (Gulati *et al.*, 2013). This initial dataset will be referred to as the 'pilot' dataset and the detailed results are deposited on a preprint server (https://doi.org/10.1101/034603). To further validate the observed pattern found in the pilot dataset, we repeated the same experiment in locally treated leaves with higher temporal resolution, including seven time points at 0.5, 1, 2, 5, 6, 11 and 13 h after simulated herbivore elicitation. This second time series dataset is the primary focus of this study and will be referred to as the 'replication' dataset when the focus is ambiguous.

The pilot dataset was obtained from the NCBI gene expression database (BioProject ID: PRJNA143589) and quantile normalization was applied to all microarray data before statistical analysis. The pilot and replication datasets contain two groups of microarray data for each tissue: an elicited group (wounding + OSs from *M. sexta* to simulate herbivore attack), and a control group (no manipulation). Both pilot and replication datasets had three biological replicates (plants) per harvest time. The statistical differences and fold-change of each gene between control (no manipulation) and the induced group (wounding + oral secretion) for each time point were calculated using the LIMMA (v.2.14) package in R (v.3.0.2). For each data point, two different transcriptome indices were calculated: the iTAI, which was calculated based on the expression fold-change and gene age, and the iTDI, which was calculated based on the expression fold-change and sequence divergence (K_a/K_s). The iTAI is a weighted mean age of the transcriptome that is induced at each time point. The iTDI is a weighted mean sequence divergence of the transcriptome that is induced at each time point. The iTAI and iTDI are analogous to the TAI and TDI found in previous studies (Kalinka *et al.*, 2010; Quint *et al.*, 2012; Cheng *et al.*, 2015; Drost *et al.*, 2015), respectively. iTAI and iTDI are defined as follows:

$$\mathrm{iTAI}_t = \frac{\sum_{i=1}^{n} \mathrm{PS}_i |\mathrm{FC}_{it}|}{\sum_{i=1}^{n} |\mathrm{FC}_{it}|} \quad \text{Eqn 1}$$

$$\mathrm{iTDI}_t = \frac{\sum_{i=1}^{n} \left(\frac{K_{ai}}{K_{si}}\right) |\mathrm{FC}_{it}|}{\sum_{i=1}^{n} |\mathrm{FC}_{it}|} \quad \text{Eqn 2}$$

where t is a time point, n is the total number of genes analyzed, PS_i is the assigned PS of gene i, $|\mathrm{FC}_{it}|$ is the absolute value of the $\log_2$ fold-change of gene i at time point t, and $\frac{K_{ai}}{K_{si}}$ is the $\frac{K_a}{K_s}$ of gene i. In order to compute these two indices, all the probes from the microarray were mapped to the predicted protein coding genes using BLASTN, and probes that could be mapped to more than one gene with 60 bp matches and 100% identity were removed.

We calculated iTAI and iTDI instead of original TAI and TDI indices for two reasons: the main objective was to analyze the evolution of induced responses, and therefore the original TAI and TDI that analyze gene expression information alone do not adequately reflect the responses induced by HAEs; and the time-course experiment lasted 21 h, and the diurnal and circadian rhythms can strongly influence gene expression changes. To minimize the effects of diurnal and circadian rhythms, we calculated induced fold-changes in OSs-treated leaves over untreated controls collected at the same time point, normalizing for any circadian/diurnal patterns.

Only genes with at least one unique probe on the microarray dataset were considered for iTAI and iTDI calculations. In total, 18 287 and 13 746 genes were analyzed from the microarray data to calculate the iTAI and iTDI, respectively.

For flg22-induced transcriptomic responses in leaves, we used expression profiles (GSE51720) based on a sequencing technique from a time-course experiment that included *A. thaliana* leaves induced by flg22 within 3 h (Rallapalli *et al.*, 2014). The $\log_2$ fold-change data generated from this dataset (Rallapalli *et al.*, 2014), and gene age and K_a/K_s information from the development hourglass study by Quint *et al.* (2012) were used for calculating iTAI and iTDI, employing the methods described earlier. For bacterial-induced transcriptomic responses, the microarray data from the AtGenExpress biotic stress dataset were used (the

data were originally downloaded from http://www.weigelworld.org/in October 2013). The $\log_2$ fold-changes were calculated using the R package LIMMA (v.2.14). The iTAI and iTDI patterns from induced transcriptome responses induced by *Phytophthora infestans*, *Pseudomonas syringae* pv. *phaseolicola* and three *P. syringae* pv. *tomato* strains (HrcC-, DC300, avrRpm1) were analyzed using the approach described earlier.

To test whether the iTAI and iTDI values were significantly different from a flat line and consistent with an hourglass pattern, the permutation tests described in Drost *et al.* (2015) were used (10 000 permutations and 100 runs). For the pilot dataset, because the iTAI and iTDI values differ among later time points, the distance values do not fit a normal distribution, which was assumed to be the case in the original method for the calculation of *P*-values (Drost *et al.*, 2015). Therefore, we used a more conservative approach for calculating *P*-values based only on three time points: the time point with the lowest iTAI or iTDI value (for herbivore dataset, 5 h) was assumed to be the middle stage, and the lowest iTAI or iTDI values before and after the middle time point were selected as the early and late stages, respectively. The *P*-values calculated using this approach are therefore usually higher (more conservative estimation) than they would be if all time points were included in the later stage samples. The conclusions about the iTAI and iTDI patterns were robust even when using this more conservative approach and when using nonparametric statistical tests on all time points. For the repeated dataset, we assigned 0.5 and 1 h as early stage time points, 2, 5, and 6 h as middle stage time points, and 11 and 13 h as late stage time points.

Log-odds ratio

We calculated the log-odds ratio of significantly induced genes within specific gene sets of interest. Rather than using the continuous fold-change data alone, which would give undue weight to genes that are strongly over- or underexpressed, classifying genes as 'induced' or 'not induced' and analyzing such groups by a log-odds ratio was more informative. The significantly induced genes were identified based on whether their expression was significantly induced by *M. sexta* OSs in comparison to untreated control (false discovery rate adjusted *P*-value < 0.05 in at least one probe and an absolute $\log_2$ fold-change > 1) for each time point. Then the log-odds ratio was calculated as:

$$\text{logodds}_{\text{gI,t}} = \log_2\left(\frac{|\mathbf{g}_{\text{I,t}}|/|\neg\mathbf{g}_{\text{I,t}}|}{|\mathbf{G}|/|\neg\mathbf{G}|}\right) \quad \text{Eqn 3}$$

where **g** is the gene set of interest, t is each time point; $|\mathbf{g}_{\text{I,t}}|$ is the number of genes that are induced (I) in **g** at time t; $|\neg\mathbf{g}_{\text{I,t}}|$ is the number of induced (I) genes at time t that are not in **g**; $|\mathbf{G}|$ is the total number of genes that belong to gene set **G** and have microarray expression data;$|\neg\mathbf{G}|$ is the total number of genes that are not in **G** and have microarray expression data. For each **G**, a log-odds ratio > 0 indicates that the genes from the tested **G** are overrepresented among the significantly OSs-induced genes, whereas a log-odds ratio < 0 indicates underrepresentation. The confidence intervals for the odds ratio were calculated by appealing to the asymptotic normality of the log-odds ratio, which has a limiting variance given by the square root of the sum of the reciprocals of $|\mathbf{G}_{\text{I,t}}|$, $|\neg\mathbf{G}_{\text{I,t}}|$, $|\mathbf{G}|$ and $|\neg\mathbf{G}|$. The gene sets analyzed in this study include two separate gene sets of primary metabolism and specialized metabolism genes, as identified in a previous study, and a gene set comprising previously identified early defense response genes (Zhou *et al.*, 2016).

Annotating genes involved in primary metabolism and the biosynthesis of specialized metabolites

We first mapped all *N. attenuata* genes to the NCBI reference sequence database (downloaded in February 2016) using BLAST with an e-value cutoff 1e–10. We then assigned enzyme commission (EC) numbers to mapped genes using BLAST2GO. Using the data provided in the supplementary files of a study by Chae *et al.* (2014), we mapped each of the *N. attenuata* EC-annotated genes to either specialized metabolism pathways or primary metabolism pathways. If a gene was found to play a role in both primary and specialized metabolism pathways, it was assigned to the primary metabolism group. The final primary metabolism gene set included 2059 genes, of which 1497 had valid expression data and 1441 were assigned a K_a/K_s. The final specialized metabolism gene set included 258 genes, of which 171 had valid expression data and 191 had a specific K_a/K_s. Differences in the average assigned PS and K_a/K_s of these gene sets relative to the entire gene set were calculated using the Wilcoxon rank-sum test.

As for the early defense response genes identified previously (Zhou *et al.*, 2016), 1274 genes in total were identified as early defense response genes, of which 1101 had valid expression data and 1056 had a specific K_a/K_s. Differences in the average assigned PS and K_a/K_s of this gene set relative to the entire gene set were also calculated using the Wilcoxon rank-sum test.

Gene ontology enrichment analysis

To further understand the mechanism of the herbivory-induced signaling hourglass pattern, gene ontology (GO) enrichment analyses were conducted on the significantly induced genes for each time point. The Cytoscape app ClueGO was used to determine significant GO groups with an adjusted *P*-value < 0.05. The REViGO online visualization tool (Supek *et al.*, 2011) was used to reduce this list of redundant GO categories and to produce TreeMaps based on the statistical significance of each category.

Available data and code

The source code and R scripts used for analyzing new microarray data are provided as Notes S1. The microarray datasets generated in this study have been submitted to NCBI under accession number GSE90951.

Results

The evolutionary age of each *N. attenuata* gene was estimated with a phylostratigraphic map, constructed by identifying the most distant phylogenetic node that contains at least one species with a detectable homolog (Kalinka *et al.*, 2010; Quint *et al.*, 2012). In total, 33 449 *N. attenuata* genes were assigned to 13 phylostratigraphic groups (phylostrata; Fig. S1), with the oldest genes assigned to phylostratum 1 (PS1, cellular organisms; shares homology with prokaryotes) and the youngest genes assigned to PS13 (*N. attenuata* specific). To estimate the sequence divergence, we calculated the K_a/K_s ratio, between *N. attenuata* and *N. obtusifolia*, which diverged *c.* 7 Ma (Sarkinen *et al.*, 2013).

To capture the evolutionary properties of genes induced by HAEs, we computed two different transcriptomic induction indices for each transcriptome after simulated herbivory: the iTAI, which is a measure of gene age (assigned PS; see the Materials and Methods section) weighted by gene induction ($\log_2$ fold-change); and the iTDI, which is a measure of sequence divergence (K_a/K_s; see the Materials and Methods section) weighted by gene induction ($\log_2$ fold-change). Here, the iTAI represents the mean evolutionary age of HAE-induced genes (phylostratum) weighted by their respective degree of induction. Similarly, iTDI represents the mean sequence divergence of HAE-induced genes, where a gene's sequence divergence (K_a/K_s) is weighted by its degree of induction (see Materials and Methods). Together, iTAI and iTDI indices provide complementary and independent measurements of evolutionary distances (Fig. S3).

We first analyzed the iTAI and iTDI using a pilot microarray dataset from a previously published, HAE-induced 21 h time-course experiment (Kim *et al.*, 2011). The gene expression was measured based on samples collected at 4 h intervals from locally treated leaves (TLs), systemic leaves (SLs) and roots (RTs), from both control and *M. sexta* OS-induced wild-type *N. attenuata* plants (Kim *et al.*, 2011). The results revealed that HAE-induced transcriptomic responses in *N. attenuata* locally treated leaves display a statistically significant evolutionary hourglass (Fig. S4): after OS elicitation, while relatively young genes with higher protein sequence divergence were induced at early and late time points, older and more evolutionarily constrained genes were induced at middle time points. A similar hourglass pattern was also found in systemic leaves but not in the roots (Fig. S5). The hourglass pattern found in systemic leaves was largely a result of a similar group of genes to those found in the TL (Table S2).

Because the middle stage of the hourglass pattern from the pilot dataset was only supported by one time point (5 h), we repeated the experiment at time points between 0.5 and 13 h with increased temporal resolution to further validate the observed hourglass pattern. The same hourglass pattern was observed in the replication experiment (Fig. 1). At 0.5 and 1 h after OS elicitation, the induced genes are relatively young (higher iTAI) and divergent (high iTDI). At 2, 5, and 6 h the induced genes are dominated by older and more constrained genes. At the late time points (11 and 13 h), the induced genes are relatively young and again divergent. The permutation tests described in Drost *et al.* (2015) revealed that the iTAI and iTDI values were significantly different from a flat line ($P = 4.1e{-}35$) and consistent with an hourglass pattern ($P = 4e{-}17$). Calculating the iTAI and iTDI for up- and down-regulated genes separately reveals that up-regulated genes are primarily responsible for the hourglass (flat-line, $P = 2.2e{-}32$; hourglass, $P = 7.8e{-}35$) (Fig. 1).

To test the robustness of the observed hourglass pattern, we compared the iTAI index based on gene ages estimated using two different homolog searching algorithms, BLASTP and PSI-BLAST. Our analyses revealed that the observed hourglass pattern was robust to different estimations of gene age and gene divergence (Fig. S2).

To further understand the mechanism underlying the observed evolutionary hourglass pattern, we performed GO enrichment analysis on the significantly induced genes at each time point (Fig. S6). At 0.5 and 1 h, the induced genes are enriched in defense and stress signal processing GO terms, such as responses to biotic stresses, defense response, JA signaling pathways, etc. These genes are known to play key roles in plant–environment interactions and are believed to be under divergent selection (Howe & Jander, 2008; Cui *et al.*, 2015). At 5 and 6 h, genes associated with GO terms of primary metabolic processes, such as oxidation–reduction process, alpha-amino acid catabolic

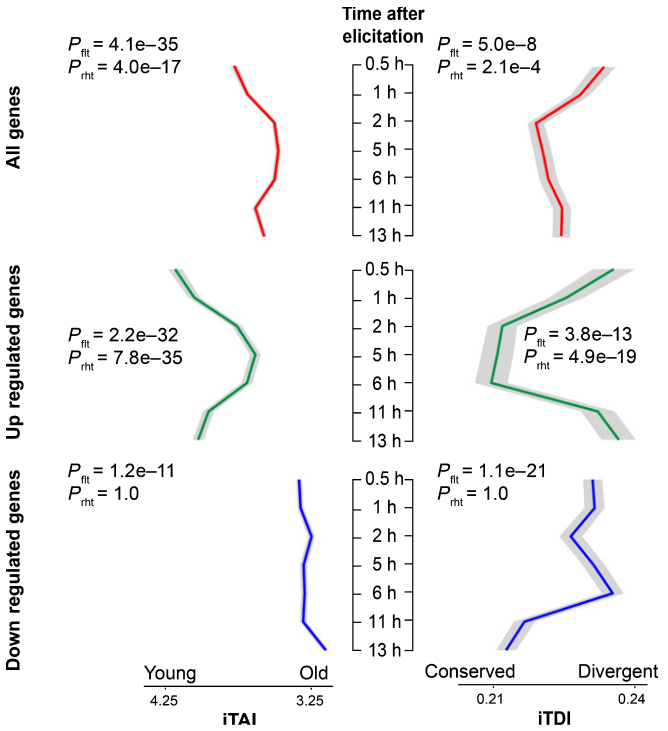

Fig. 1 Transcriptome induction age index (iTAI) and transcriptome induction divergence index (iTDI) in locally treated *Nicotiana attenuata* leaves after simulated herbivory are shown in left and right, respectively. Red, green and blue lines designate mean indices calculated on all genes, up- and down-regulated genes, respectively. The gray ribbons refer to SD. P_{flt} and P_{rht} indicate the *P*-values from a flat line and reductive hourglass tests, respectively. $P_{flat} < 0.05$ indicates that the pattern is significantly different from a flat line and $P_{rht} < 0.05$ indicates that it follows an hourglass (high–low–high) pattern. Time points were assigned to three different stages: early (0.5 and 1 h), middle (2, 5 and 6 h) and late (11 and 13 h).

process, were highly enriched. At 11 and 13 h, genes associated with secondary metabolism and defense responses were enriched. Consistently, several genes involved in the biosynthesis of *Nicotiana* or Solanaceae-specific defense metabolites, such as HGL-DTGs, which are potent herbivore toxins (Heiling *et al.*, 2010), were among the highest induced genes at 11 and 13 h. These results indicate that the observed evolutionary hourglass results from the architecture and modulation of an HAE-induced signaling cascade in *N. attenuata* leaves.

To provide further statistical analysis for the modulation of genes induced at different stages, we assigned genes to three groups: early defense signaling genes, which were previously identified based on a transcriptomic network analysis ($n = 1101$), genes involved in primary metabolism ($n = 1497$), and genes involved in the biosynthesis of specialized metabolites ($n = 171$), which were identified using enzyme annotations from a previous study (Chae *et al.*, 2014). Consistent with our prediction, primary metabolism genes are much older compared with the genome-wide background (average PS for primary metabolism genes = 2.55; mean PS all genes = 4.71; $P < 2.2e{-}16$), genes involved in early defense signaling (average PS for early defense genes = 4.77; $P < 2.2e{-}16$), and genes involved in the biosynthesis of specialized metabolites (average PS for specialized metabolism genes = 4.31; $P < 2.2e{-}16$). Similarly, the genes involved in primary metabolism are more conserved than the genome-wide average (average K_a/K_s for primary metabolism genes = 0.17; average K_a/K_s all genes = 0.26; $P < 2.2e{-}16$), early defense response genes (average K_a/K_s for early defense response genes = 0.21; $P = 8.9e{-}15$), and genes involved in specialized metabolism (average K_a/K_s for specialized metabolism genes 0.19; $P = 0.02$). Because many genes that are known to be involved in the biosynthesis of *Nicotiana* or Solanaceae-specific specialized metabolites, such as HGL-DTGs, which are potent herbivore toxins (Heiling *et al.*, 2010), were not annotated by Chae *et al.* (2014), we further analyzed the K_a/K_s ratio of three genes involved in the late steps of HGL-DTG biosynthesis that were highly induced at 11 and 13 h. On average, these three genes have a K_a/K_s ratio of 0.33 ± 0.11, much higher than the genome-wide average (0.23) and genes involved in primary metabolism (0.16).

The log-odd ratio analysis showed that after OS elicitation, the early defense signaling genes were most enriched at early time points (0.5 and 1 h) (Fig. 2a), while genes involved in primary metabolism and biosynthesis of specialized metabolites were most enriched at middle (5 and 6 h) and late (11 and 13 h) time points (Fig. 2b,c). This pattern is consistent with the hypothesis that the evolutionary hourglass pattern of the herbivory-induced defenses reflects divergent responses to selection among different modules of the induced signaling cascade (Fig. 3).

Is the observed hourglass pattern specific to HAE-induced defense signaling? We computed iTAI and iTDI values from locally treated leaves of *A. thaliana* induced by flg22, a bacterial-associated elicitor (Rallapalli *et al.*, 2014). Throughout the 3 h time course of this experiment, both iTAI and iTDI decreased (Fig. S7), consistent with the HAE-induced pattern observed at early time points (0.5–6 h) in *N. attenuata*. As data are not available for later times in the elicitation process, we do not know if

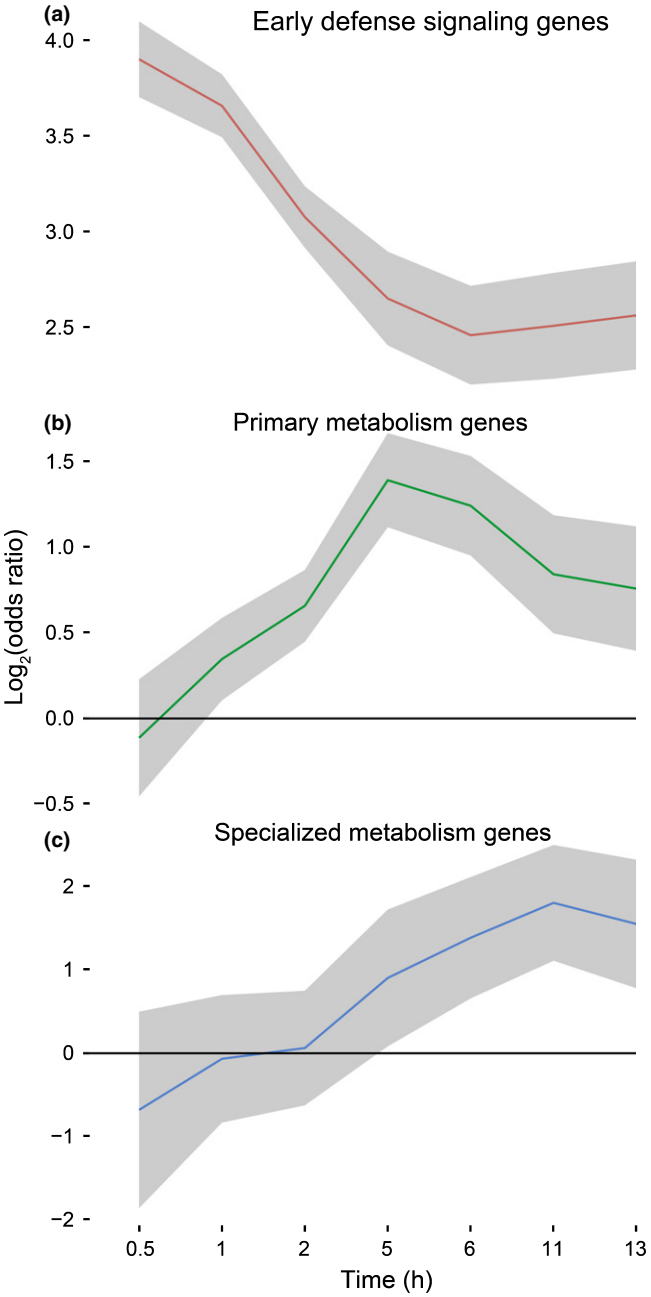

Fig. 2 Log-odds ratio of herbivory-induced genes involved in early defense signaling, primary metabolisms, and specialized metabolisms in *Nicotiana attenuata*. (a) The previously identified early defense signaling genes are overrepresented most at early time points following simulated herbivore attack; (b) genes in primary metabolism are more enriched at middle time points than at early and late time points. (c) Specialized metabolism genes are overrepresented primarily at late time points (11 and 13 h). The gray ribbons represent the 95% confidence interval. Genes that showed significant up-regulation (false discovery rate adjusted *P*-value < 0.05 and fold-change > 2) by simulated herbivory were considered.

flg22 elicits the exact same hourglass pattern as HAE elicitation. However, the early pattern found for flg22-induced iTAI and iTDI indices is consistent with a general hourglass response pattern in plant biotic stress-induced defense signaling. Interestingly, the iTAI and iTDI showed different patterns when plants were

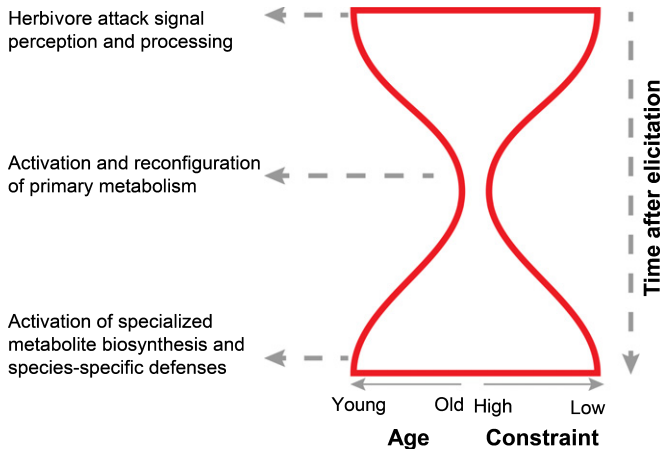

Fig. 3 Evolutionary hourglass of herbivory-induced defense signaling in *Nicotiana attenuata*. Herbivory-induced defense signaling is modulated in three phases. (1) Immediately after a plant perceives herbivore attack, a large group of rapidly evolving and relatively young genes involved in elicitor perception and processing are up-regulated; (2) the perceived and processed signals result in up-regulation of highly conserved and ancient genes involved in primary metabolism; and (3) at later time points, the reprogrammed transcriptomes activate young and rapidly evolving genes involved in defenses and the biosynthesis of specialized metabolites.

infected by different living pathogens (Fig. S8), either because of low temporal resolution of the time series, or perhaps because of changes in defense signaling caused by pathogen–plant interactions that follow the early pathogen recognition responses.

Discussion

Our results show that the distinct gene modules involved in herbivory-induced defense responses have evolved differently in plants (Fig. 3) and display an hourglass pattern. After herbivory attack, the genes that are induced at the middle stage (5–6 h) were much older (lower PS) and conserved (lower K_a/K_s) than those genes that are induced at early (0.5–2 h) and late stages (9–21 h).

When analyzing both a previously published time-course dataset (pilot) and a newly collected dataset (replication) with greater temporal resolution, we found the same evolutionary hourglass pattern, indicating the robustness of the findings. The only difference between these two datasets was the genes that were induced at 5 h after elicitation. In the pilot dataset, genes involved in RNA translation and modifications were highly enriched (Fig. S9), while in the replication dataset, genes involved in the biosynthesis and regulation of primary and specialized metabolites were enriched at this same time point. As a large proportion of genes that were induced at 5 h were also found at 2 and 6 h after elicitation in the replication dataset (Figs 2, S6), we reasoned that the induction of genes involved in RNA translation and modifications from the pilot dataset might represent only a short temporal phase, which might be sensitive to experimental conditions.

The observed hourglass pattern represents the early episodes of herbivory-induced transcriptomic responses. Clearly genes are still induced later, beyond 21 h after the elicitation; however, it is likely that the total number of genes that are induced is lower. From both the pilot and replicate datasets, the total number of significantly induced genes decreases continuously with time after elicitation. For example, the numbers of genes that were significantly induced at the last time point (21 or 13 h) are only 23.3% and 33% of the number of genes that were significantly induced at 1 h in the pilot and replicate datasets, respectively. We therefore speculate that the magnitude of transcriptomic responses after 21 h would be much less than in the first 21 h. Future studies that expand our analysis to capture longer temporal dynamics of herbivory-induced transcriptomic responses will shed more light on whether the observed hourglass pattern is specific to the early episode of induced responses or is part of an oscillation pattern.

Although lacking data at later time points, both the iTAI and iTDI patterns of the flg22-induced transcriptomic responses in *Arabidopsis* leaves are consistent with the herbivory-induced defense responses in *N. attenuata*, suggesting that the evolutionary hourglass might represent a general model for describing signaling cascades that are induced by biotic stresses. Clearly, many more resistance responses elicited by different biotic and abiotic stimuli need to be examined with similar phylotranscriptomic approaches in different plant species to test the robustness of the hourglass phenomena as a general model for signaling cascade evolution. The challenge will be to develop/find/mine datasets that are sufficiently deep in their temporal resolution to capture the complete breadth of a response, without the pattern being confounded by additional cycles of elicitation and response, as commonly occurs in biotrophic pathogen–plant interactions.

The evolutionary pattern of herbivory-induced transcriptomic responses resembles the hourglass pattern that is often used to describe embryogenesis in animals (Domazet-Loso & Tautz, 2010; Kalinka *et al.*, 2010) and plants (Quint *et al.*, 2012), as well as general developmental processes in plants and fungi (Cheng *et al.*, 2015; Drost *et al.*, 2016). This is probably a result of similar underlying mechanisms that led to the evolutionary hourglass pattern. For animal embryogenesis, two nonexclusive hypotheses are proposed to explain the observed hourglass pattern. First, it is thought that the conserved phylotypic period (the middle stage of animal embryogenesis) reflects the evolutionary constraints of body plan for animals, as the HOX gene cluster was found to be highly expressed during this stage (Kalinka *et al.*, 2010; Irie & Kuratani, 2014). Second, it is hypothesized that the hourglass pattern reflects differential interactions with ecological factors at different developmental stages. While early (zygote) and late (juvenile and adults) embryonic stages often interact with environmental stimuli, the mid-embryonic stages that characterize the phylotypic phase are normally not in direct contact with the environment and are thus less likely to be subject to ecological adaptations and evolutionary changes (Domazet-Loso & Tautz, 2010). For herbivory-induced defenses, the hourglass pattern is likely to be a result of either strong evolutionary constraints on the genes involved in primary metabolism that were activated at the middle stage (hypothesis 1), or high divergent selection or relaxed purifying selection on genes that were

involved in early defense signaling and specialized metabolism (hypothesis 2). These two hypotheses are not mutually exclusive and both are supported by genome-wide analysis or by individual gene cases. On the one hand, herbivory-induced alterations in primary metabolism have been found in many plant systems (Zhou *et al.*, 2015), and are receiving increased attention (Machado *et al.*, 2013, 2015; Robert *et al.*, 2014; Zhou *et al.*, 2015). Genes involved in primary metabolism evolved much earlier and are thought to be under stronger evolutionary constraints than are those of specialized metabolism in plants (Weng, 2014; Chae *et al.*, 2015; Mukherjee *et al.*, 2015), probably as a result of more pleiotropic effects, consistent with hypothesis 1. On the other hand, genes involved in early defense signaling and specialized metabolism have been shown to be under positive or relaxed purifying selections (Yang *et al.*, 2002; Tiffin & Moeller, 2006; Howe & Jander, 2008; Chen *et al.*, 2011; Cui *et al.*, 2015), consistent with hypothesis 2.

Systems biologists have predicted that the modularization and hourglass architecture (bow-tie shape) of signaling networks can facilitate both stability against perturbations (robustness) and evolvability of traits (Kitano, 2004; Friedlander *et al.*, 2015); both of these features are required for developmental processes and induced defense. Therefore, it is plausible that the modulation of the signaling network itself in induced defenses and embryogenesis might be a consequence of adaptations that facilitated their robustness. Synthetic allopolyploid plants may represent an excellent system in which to test this inference. Allopolyploidy, whether laboratory-produced or natural, occurs when the genomes of two different species fuse and new signaling systems are produced that emerge from the recruitment of different modules of the parental species in new combinations (Anssour & Baldwin, 2010). A prediction arising from this study is that the variation of induced defenses among offspring produced from these interspecies fusions will display a similar hourglass pattern, in which transcriptomic responses at early and late time points after elicitation are more variable than the ones in the middle.

Acknowledgements

We thank T. Brockmüller and Z. Ling for help with data analysis and discussions, C. Kreitzer for collecting samples, W. Kröber for microarray analysis and K. Gase for submitting the data to the NCBI. We are grateful for the funding by the Swiss National Science Foundation (project no. PEBZP3-142886 to S.X.), the Marie Curie Intra-European Fellowship (IEF, project no. 328935 to S.X.), the Max Planck Society, European Research Council advanced grant ClockworkGreen (project no. 293926 to I.T.B.), and Brigham Young University (travel grants to J.B. and M.D.).

Author contributions

M.D. and S.X. planned the research, and M.D., J.B., and S.X. performed data analysis. W.Z. performed the new kinetic experiment. M.D., J.B., I.T.B., and S.X. wrote the manuscript.

References

Agrawal AA. 1998. Induced responses to herbivory and increased plant performance. *Science* **279**: 1201–1202.

Agrawal AA, Conner JK, Johnson MTJ, Wallsgrove R. 2002. Ecological genetics of an induced plant defense against herbivores: additive genetic variance and costs of phenotypic plasticity. *Evolution* **56**: 2206–2213.

Anssour S, Baldwin IT. 2010. Variation in antiherbivore defense responses in synthetic *Nicotiana* allopolyploids correlates with changes in uniparental patterns of gene expression. *Plant Physiology* **153**: 1907–1918.

Baldwin IT. 1998. Jasmonate-induced responses are costly but benefit plants under attack in native populations. *Proceedings of the National Academy of Sciences, USA* **95**: 8113–8118.

Baldwin IT, Hamilton W. 2000. Jasmonate-induced responses of *Nicotiana sylvestris* results in fitness costs due to impaired competitive ability for nitrogen. *Journal of Chemical Ecology* **26**: 915–952.

Bodenhausen N, Reymond P. 2007. Signaling pathways controlling induced resistance to insect herbivores in *Arabidopsis*. *Molecular Plant-Microbe Interactions* **20**: 1406–1420.

Bonaventure G, VanDoorn A, Baldwin IT. 2011. Herbivore-associated elicitors: FAC signaling and metabolism. *Trends in Plant Science* **16**: 294–299.

Chae L, Kim T, Nilo-Poyanco R, Rhee SY. 2014. Genomic signatures of specialized metabolism in plants. *Science* **344**: 510–513.

Chae L, Kim T, Nilo-Poyanco R, Rhee SY. 2015. Genomic signatures of specialized metabolism in plants. *Science* **344**: 510–513.

Chen F, Tholl D, Bohlmann J, Pichersky E. 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant Journal* **66**: 212–229.

Cheng XJ, Hui JHL, Lee YY, Law PTW, Kwan HS. 2015. A "developmental hourglass" in fungi. *Molecular Biology and Evolution* **32**: 1556–1566.

Cui HT, Tsuda K, Parker JE. 2015. Effector-triggered immunity: from pathogen perception to robust defense. *Annual Review of Plant Biology* **66**: 487–511.

van Dam NM, Baldwin IT. 1998. Costs of jasmonate-induced responses in plants competing for limited resources. *Ecology Letters* **1**: 30–33.

Domazet-Loso T, Brajkovic J, Tautz D. 2007. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends in Genetics* **23**: 533–539.

Domazet-Loso T, Tautz D. 2010. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* **468**: 815–818.

Drost HG, Bellstadt J, O'Maoileidigh DS, Silva AT, Gabel A, Weinholdt C, Ryan PT, Dekkers BJ, Bentsink L, Hilhorst HW *et al.* 2016. Post-embryonic hourglass patterns mark ontogenetic transitions in plant development. *Molecular Biology and Evolution* **33**: 1158–1163.

Drost HG, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. *Molecular Biology and Evolution* **32**: 1221–1231.

Erb M, Veyrat N, Robert CAM, Xu H, Frey M, Ton J, Turlings TCJ. 2015. Indole is an essential herbivore-induced volatile priming signal in maize. *Nature Communications* **6**: 6273.

Friedlander T, Mayo AE, Tlusty T, Alon U. 2015. Evolution of bow-tie architectures in biology. *PLoS Computational Biology* **11**: e1004055.

Gaquerel E, Gulati J, Baldwin IT. 2014. Revealing insect herbivory-induced phenolamide metabolism: from single genes to metabolic network plasticity analysis. *Plant Journal* **79**: 679–692.

Glawe GA, Zavala JA, Kessler A, Van Dam NM, Baldwin IT. 2003. Ecological costs and benefits correlated with trypsin protease inhibitor production in *Nicotiana attenuata*. *Ecology* **84**: 79–90.

Gulati J, Kim SG, Baldwin IT, Gaquerel E. 2013. Deciphering herbivory-induced gene-to-metabolite dynamics in *Nicotiana attenuata* tissues using a multifactorial approach. *Plant Physiology* **162**: 1042–1059.

Heil M, Baldwin IT. 2002. Fitness costs of induced resistance: emerging experimental support for a slippery concept. *Trends in Plant Science* **7**: 61–67.

Heiling S, Schuman MC, Schoettner M, Mukerjee P, Berger B, Schneider B, Jassbi AR, Baldwin IT. 2010. Jasmonate and ppHsystemin regulate key malonylation steps in the biosynthesis of 17-hydroxygeranyllinalool diterpene

glycosides, an abundant and effective direct defense against herbivores in *Nicotiana attenuata*. *Plant Cell* **22**: 273–292.

Howe GA, Jander G. 2008. Plant immunity to insect herbivores. *Annual Review of Plant Biology* **59**: 41–66.

Irie N, Kuratani S. 2014. The developmental hourglass model: a predictor of the basic body plan? *Development* **141**: 4649–4655.

Jones DT, Swindells MB. 2002. Getting the most from PSI-BLAST. *Trends in Biochemical Sciences* **27**: 161–164.

Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, Ohler U, Bergman CM, Tomancak P. 2010. Gene expression divergence recapitulates the developmental hourglass model. *Nature* **468**: 811–814.

Karban R, Baldwin IT. 1997. *Induced responses to herbivory*. Chicago, IL, USA: University of Chicago Press.

Kaur H, Heinzel N, Schottner M, Baldwin IT, Galis I. 2010. R2R3-NaMYB8 regulates the accumulation of phenylpropanoid-polyamine conjugates, which are essential for local and systemic defense against insect herbivores in *Nicotiana attenuata*. *Plant Physiology* **152**: 1731–1747.

Kessler A, Halitschke R, Baldwin IT. 2004. Silencing the jasmonate cascade: induced plant defenses and insect populations. *Science* **305**: 665–668.

Kim SG, Yon F, Gaquerel E, Gulati J, Baldwin IT. 2011. Tissue specific diurnal rhythms of metabolites and their regulation during herbivore attack in a native tobacco, *Nicotiana attenuata*. *PLoS ONE* **6**: e26214.

Kitano H. 2004. Biological robustness. *Nature Reviews Genetics* **5**: 826–837.

Krügel T, Lim M, Gase K, Halitschke R, Baldwin IT. 2002. *Agrobacterium*-mediated transformation of *Nicotiana attenuata*, a model ecological expression system. *Chemoecology* **12**: 177–183.

Li D, Baldwin IT, Gaquerel E. 2015a. Navigating natural variation in herbivory-induced secondary metabolism in coyote tobacco populations using MS/MS structural analysis. *Proceedings of the National Academy of Sciences, USA* **112**: E4147–E4155.

Li R, Zhang J, Li JC, Zhou GX, Wang Q, Bian WB, Erb M, Lou YG. 2015b. Prioritizing plant defence over growth through WRKY regulation facilitates infestation by non-target herbivores. *eLife* **4**: e04805.

Liu QS, Wang XY, Tzin V, Romeis J, Peng YF, Li YH. 2016. Combined transcriptome and metabolome analyses to understand the dynamic responses of rice plants to attack by the rice stem borer *Chilo suppressalis* (Lepidoptera: Crambidae). *BMC Plant Biology* **16**: 259.

Machado RA, Arce CC, Ferrieri AP, Baldwin IT, Erb M. 2015. Jasmonate-dependent depletion of soluble sugars compromises plant resistance to *Manduca sexta*. *New Phytologist* **207**: 91–105.

Machado RAR, Ferrieri AP, Robert CAM, Glauser G, Kallenbach M, Baldwin IT, Erb M. 2013. Leaf-herbivore attack reduces carbon reserves and regrowth from the roots via jasmonate and auxin signaling. *New Phytologist* **200**: 1234–1246.

Mattiacci L, Dicke M, Posthumus MA. 1995. beta-Glucosidase – an elicitor of herbivore-induced plant odor that attracts host-searching parasitic wasps. *Proceedings of the National Academy of Sciences, USA* **92**: 2036–2040.

Moyers BA, Zhang JZ. 2015. Phylostratigraphic bias creates spurious patterns of genome evolution. *Molecular Biology and Evolution* **32**: 258–267.

Mukherjee D, Mukherjee A, Ghosh TC. 2015. Evolutionary rate heterogeneity of primary and secondary metabolic pathway genes in *Arabidopsis thaliana*. *Genome Biology and Evolution* **8**: 17–28.

Qi JF, Sun GL, Wang L, Zhao CX, Hettenhausen C, Schuman MC, Baldwin IT, Li J, Song J, Liu ZD *et al.* 2016. Oral secretions from *Mythimna separata* insects specifically induce defence responses in maize as revealed by high-dimensional biological data. *Plant, Cell & Environment* **39**: 1749–1766.

Quint M, Drost HG, Gabel A, Ullrich KK, Bonn M, Grosse I. 2012. A transcriptomic hourglass in plant embryogenesis. *Nature* **490**: 98–101.

Rallapalli G, Kemen EM, Robert-Seilaniantz A, Segonzac C, Etherington GJ, Sohn KH, MacLean D, Jones JDG. 2014. EXPRSS: an Illumina based high-throughput expression-profiling method to reveal transcriptional dynamics. *BMC Genomics* **15**: 341.

Redman AM, Cipollini DF, Schultz JC. 2001. Fitness costs of jasmonic acid-induced defense in tomato, *Lycopersicon esculentum*. *Oecologia* **126**: 380–385.

Robert CAM, Ferrieri RA, Schirmer S, Babst BA, Schueller MJ, Machado RAR, Arce CCM, Hibbard BE, Gershenzon J, Turlings TCJ *et al.* 2014. Induced carbon reallocation and compensatory growth as root herbivore tolerance mechanisms. *Plant, Cell & Environment* **37**: 2613–2622.

Sarkinen T, Bohs L, Olmstead RG, Knapp S. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology* **13**: 214.

Strauss SY, Rudgers JA, Lau JA, Irwin RE. 2002. Direct and ecological costs of resistance to herbivory. *Trends in Ecology & Evolution* **17**: 278–285.

Supek F, Bosnjak M, Skunca N, Smuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6**: e21800.

Tiffin P, Moeller DA. 2006. Molecular evolution of plant immune system genes. *Trends in Genetics* **22**: 662–670.

Turlings TCJ, Tumlinson JH, Lewis WJ. 1990. Exploitation of herbivore-induced plant odors by host-seeking parasitic wasps. *Science* **250**: 1251–1253.

Tzin V, Fernandez-Pozo N, Richter A, Schmelz EA, Schoettner M, Schafer M, Ahern KR, Meihls LN, Kaur H, Huffaker A *et al.* 2015. Dynamic maize responses to aphid feeding are revealed by a time series of transcriptomic and metabolomic assays. *Plant Physiology* **169**: 1727–1743.

Unsicker SB, Kunert G, Gershenzon J. 2009. Protective perfumes: the role of vegetative volatiles in plant defense against herbivores. *Current Opinion in Plant Biology* **12**: 479–485.

Wang XD, Wang Y, Ou LJ, He XJ, Chen D. 2015. Allocation costs associated with induced defense in *Phaeocystis globosa* (Prymnesiophyceae): the effects of nutrient availability. *Scientific Reports* **5**: 10850.

Weng JK. 2014. The evolutionary paths towards complexity: a metabolic perspective. *New Phytologist* **201**: 1141–1149.

Wu J, Baldwin IT. 2010. New insights into plant responses to the attack from insect herbivores. *Annual Review of Genetics* **44**: 1–24.

Xu S, Zhou WW, Pottinger S, Baldwin IT. 2015. Herbivore associated elicitor-induced defences are highly specific among closely related *Nicotiana* species. *BMC Plant Biology* **15**: 2.

Yang J, Huang J, Gu H, Zhong Y, Yang Z. 2002. Duplication and adaptive evolution of the chalcone synthase genes of *Dendranthema* (Asteraceae). *Molecular Biology and Evolution* **19**: 1752–1759.

Zhou W, Brockmöller T, Ling Z, Omdahl A, Baldwin IT, Xu S. 2016. Evolution of herbivore-induced early defense signaling was shaped by genome-wide duplications in Nicotiana. *eLife* **5**: e19531.

Zhou S, Lou YR, Tzin V, Jander G. 2015. Alteration of plant primary metabolism in response to insect herbivory. *Plant Physiology* **169**: 1488–1498.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Relative age assignments of genes using BLASTP and PSI-BLAST.

Fig. S2 iTAI calculated using BLASTP and PSI-BLAST.

Fig. S3 Analysis of phylostratum and K_a/K_s shows a weak correlation between them, suggesting that the two metrics are independent.

Fig. S4 Hourglass pattern found through analysis of the publicly available, 'pilot' dataset.

Fig. S5 Herbivore attack induced iTDI and iTAI in systemic leaves and roots.

Fig. S6 GO enrichment TreeMaps separated by time point.

Fig. S7 iTAI and iTDI in flg22-induced *Arabidopsis thaliana* leaves.

Fig. S8 iTAI and iTDI following introduction of five different pathogens introduced to *Arabidopsis thaliana* leaves.

Fig. S9 GO enrichment TreeMaps from the pilot study data.

Table S1 The correlation between iTAI and iTDI values in roots, treated leaves, and systemic leaves

Table S2 Percentage of up-regulated genes in systemic leaves also up-regulated in locally treated leaves

Notes S1 The source code and R scripts used for analyzing new microarray data.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.