

PLANT SCIENCES

Information theory tests critical predictions of plant defense theory for specialized metabolism

Dapeng Li¹, Rayko Halitschke¹, Ian T. Baldwin^{1*}, Emmanuel Gaquerel^{2,3*}

Different plant defense theories have provided important theoretical guidance in explaining patterns in plant specialized metabolism, but their critical predictions remain to be tested. Here, we systematically explored the metabolomes of *Nicotiana attenuata*, from single plants to populations, as well as of closely related species, using unbiased tandem mass spectrometry (MS/MS) analyses and processed the abundances of compound spectrum-based MS features within an information theory framework to test critical predictions of optimal defense (OD) and moving target (MT) theories. Information components of plant metabolomes were consistent with the OD theory but contradicted the main prediction of the MT theory for herbivory-induced dynamics of metabolome compositions. From micro- to macroevolutionary scales, jasmonate signaling was confirmed as the master determinant of OD, while ethylene signaling provided fine-tuning for herbivore-specific responses annotated via MS/MS molecular networks.

INTRODUCTION

Structurally diverse specialized metabolites are central players in plants' adaptations to their environments and particularly in their defense against enemies (1). The spectacular diversification of specialized metabolism found in plants inspired several decades of intense research about its multifaceted ecological functions and nucleated a long list of plant defense theories that provided important guidance to empirical studies of the evolution and ecology of plant-insect interactions (2). However, these plant defense theories did not follow the canonical path of hypothetical-deductive reasoning in which critical predictions are posed at the same levels of analysis (3) and tested experimentally to advance the next cycle of theory development (4). Technical constraints limited data acquisition to specific metabolic classes, precluded the comprehensive profiling of specialized metabolites, and consequently prevented the among-taxa comparisons that were essential for theory advancement (5). This lack of comprehensive metabolomics data and the processing workflows to compare the metabolic space among different plant taxa in a common currency thwarted the scientific maturation of the field.

Recent advances in the field of tandem mass spectrometry (MS/MS) metabolomics are allowing for comprehensive characterizations of metabolic variations within and among species of given phylogenetic clades and can be combined with computational methods for the calculation of structural similarity among these complex mixtures without a priori chemical knowledge (5). This combination of analytical and computational advances provides the missing framework required for the rigorous testing of many predictions made by the long-standing ecological and evolutionary theories on metabolic diversity. Information theory, first introduced by Shannon (6) in his seminal article in 1948, laid the groundwork for a mathematical analysis of information, which has been adopted by many fields beyond its original applications. In genomics, information theory has been successfully applied to quantify sequence conservation

information (7), and in a transcriptomics study, information theory parsed the global changes that occur in the transcriptome (8). In a previous study, we applied the information theory statistical framework to metabolomics to describe tissue-level metabolic specialization in plants (9). Here, we combine an MS/MS-based workflow with the statistical framework of information theory to characterize metabolic diversity in a common currency so as to compare critical predictions of plant defense theories for herbivory-induced metabolomes.

Plant defense theoretical frameworks are frequently mutually inclusive and can be classified into two groups: those that attempt to explain the distribution of plant specialized metabolites based on defensive function, such as the optimal defense (OD) (10), moving target (MT) (11), and apparency (12) theories, whereas others seek mechanistic explanations in how variation in resource availability influences plant growth and specialized metabolite accumulations, for instance, the carbon:nutrient balance hypothesis (13), the growth rate hypothesis (14), and the growth-differentiation balance hypothesis (15). These two groups of theories are posed at different levels of analysis (4). However, two theories, both addressing defensive functions at the functional level, have dominated the dialogue about plant constitutive and inducible defenses: the OD theory, which hypothesized that plants invest their costly chemical defense only when needed, for instance, when attacked by herbivores, thus directionally allocating compounds with defensive function according to the probability of future attack (10); and the MT hypothesis, which proposes that axes of directional metabolite changes do not exist but rather that metabolites change randomly, thereby creating a metabolic "moving target" that could thwart the performance of attacking herbivores. That is, the two theories present contrasting predictions about the metabolic reconfigurations that occur after herbivore attack: the unidirectional accumulations of metabolites with defense functions (OD) versus nondirectional metabolic changes (MT) (11).

The OD and MT hypotheses address not only induced changes within metabolomes but also the ecological and evolutionary consequences of these metabolite accumulations, such as the associated fitness costs and benefits of these metabolic changes in specific ecological contexts (16). While both hypotheses acknowledge the

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

¹Max Planck Institute for Chemical Ecology, Department of Molecular Ecology, Jena, Germany. ²Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, Strasbourg, France. ³Centre for Organismal Studies, University of Heidelberg, Heidelberg, Germany.

*Corresponding author. Email: emmanuel.gaquerel@ibmp-cnrs.unistra.fr (E.G.); baldwin@ice.mpg.de (I.T.B.)

defensive function of specialized metabolites, which may or may not be costly to produce, critical predictions that differentiate OD and MT hypotheses lie in the directionality of induced metabolic changes. The predictions of the OD theory have received the most experimental attention to date. These tests include investigations of the variations in specific compound classes with either direct or indirect defense functions across different tissues and ontogenetic stages in plants under both glasshouse and natural conditions (17–19). However, to date, the central distinguishing prediction of the two theories, namely, the directionality of the metabolic changes, remains to be tested because of the lack of workflows and statistical frameworks to conduct a globally comprehensive analysis of metabolic diversity of any organism. Here, we provide such an analysis.

One of the most notable characteristics of plant specialized metabolites is their extreme structural diversity at all levels ranging from single plants, populations to congeneric species (20). Much of the quantitative variation in specialized metabolites can be observed at a population scale, whereas strong qualitative differences are commonly maintained at the level of species (20). Plant metabolic diversity is hence a primary dimension of functional diversity, reflecting adaptations to different ecological niches, particularly those that differ in the probability of attack from a broad range of insects, including specialist and generalist herbivores (21). Interactions with various insects have been recognized as important selective pressures since Fraenkel's (22) seminal article on the *raison d'être* of plant specialized metabolites, and these interactions are thought to have sculpted plant metabolic pathways during evolution (23). Interspecies variations in specialized metabolite diversity may additionally reflect physiological trade-offs associated with constitutive and induced plant defense strategies against herbivory, as the two are frequently negatively correlated with each other across species (24). While being well defended at all times may be advantageous, timely defense-related metabolic changes provide clear benefits in allowing plants to allocate valuable resources to other physiological investments (19, 24) and avoid collateral damage to mutualists (25). In addition, these insect herbivory-induced reconfigurations of specialized metabolite production can lead to disruptive distributions in populations (26) and may reflect direct readouts of substantial natural variation in jasmonate (JA) signaling, with high and low JA signaling being likely maintained in populations by trade-offs between defense against herbivores and competition with conspecifics (27). Furthermore, specialized metabolite biosynthetic pathways can undergo rapid loss and gain transitions during evolution, resulting in patchy metabolic distributions among closely related species (28). These polymorphisms can be rapidly established in response to changing herbivory regimes (29), implying that fluctuations in herbivore communities are key factors driving metabolic heterogeneity.

Here, we specifically addressed the following questions. (i) How are plant metabolomes reconfigured by insect herbivory? (ii) What principle information components of the metabolic plasticity can be quantified to test predictions of long-standing defense theories? (iii) Is the reprogramming of plant metabolomes conducted in an attacker-specific fashion, and if so, what roles do phytohormones play in tailoring the specific metabolic responses, and which metabolites contribute to the species specificity of elicited defense? (iv) As many defense theories make predictions that scale across levels of biological organization, we asked how consistently do the elicited metabolic responses scale from intra- to interspecific comparisons? To this end, we systematically explored the leaf metabolomes of

Nicotiana attenuata, an ecological model plant with rich specialized metabolism and its response to attack from two native herbivores, larvae of lepidopteran *Manduca sexta* (Ms), a tobacco hornworm “specialist” that feeds largely on Solanaceae plants and *Spodoptera littoralis* (Sl), the cotton leaf worm, a “generalist” that feeds on host plants from the Solanaceae plus a broad range of other host plants from other genera and families. We parsed MS/MS metabolomics spectra and extracted information theory statistical descriptors to contrast predictions of OD and MT theories. Specificity maps were created to reveal the identity of key metabolites. The analysis was extended to *N. attenuata*'s native populations and closely related *Nicotiana* species to further dissect covariations among phytohormone signaling and OD inductions.

RESULTS

Statistical descriptors of metabolic plasticity and information content from large-scale MS/MS data

To capture a holistic picture of the plasticity and structure of the herbivory-induced leaf metabolomes of *N. attenuata*, we used a previously developed analytical and computational workflow to comprehensively collect and deconvolute high-resolution data-independent MS/MS spectra from plant extracts (9). This indiscriminant approach (termed MS/MS) allows for the construction of nonredundant compound spectra that are subsequently used for all compound-level analyses presented here. These deconvoluted plant metabolomes are highly diverse and composed of hundreds to thousands of metabolites (here, ~500- to 1000-s MS/MSs). Here, we consider metabolic plasticity in an information theory framework and quantify the diversity and specialization of a metabolome based on the Shannon entropy of the metabolic frequency distribution. Using previously implemented formulae (8), we calculated a set of indices that allow for the quantification of metabolome diversity (Hj index), metabolic profile specialization (δ_j index), and metabolic specificity of individual metabolites (Si index). In addition, we applied a relative distance plasticity index (RDPI) to quantify metabolome inducibility by herbivory (Fig. 1A) (30). Within this statistical framework, we consider MS/MS spectra as basic information units and process MS/MSs relative abundances into frequency profiles from which metabolome diversity is estimated using Shannon entropy. Metabolome specialization is measured as the average degree of specificity of individual MS/MS spectra. Consequently, an increase in abundance of certain groups of MS/MSs after herbivore elicitation translate into an increase in profile inducibility, the RDPI, and specialization, the δ_j index, as more specialized metabolites are produced, generating high Si indices, whereas a decrease in Hj diversity index reflects that either fewer MS/MSs are being produced or that the profile frequency distribution is changing toward less uniformity, simultaneously reducing its overall uncertainty. With the Si index calculation, it was possible to highlight which MS/MSs are specific to certain herbivory elicitation and, reciprocally, which are relatively nonresponsive to elicitation, a key metric that allows MT and OD predictions to be distinguished.

Predictions of plant defense theories reformulated in the axes of information theory descriptors

Using information theory descriptors, we interpret OD theory to predict that the herbivory-induced changes in specialized metabolites from the uninduced constitutive state will result in (i) an increase in

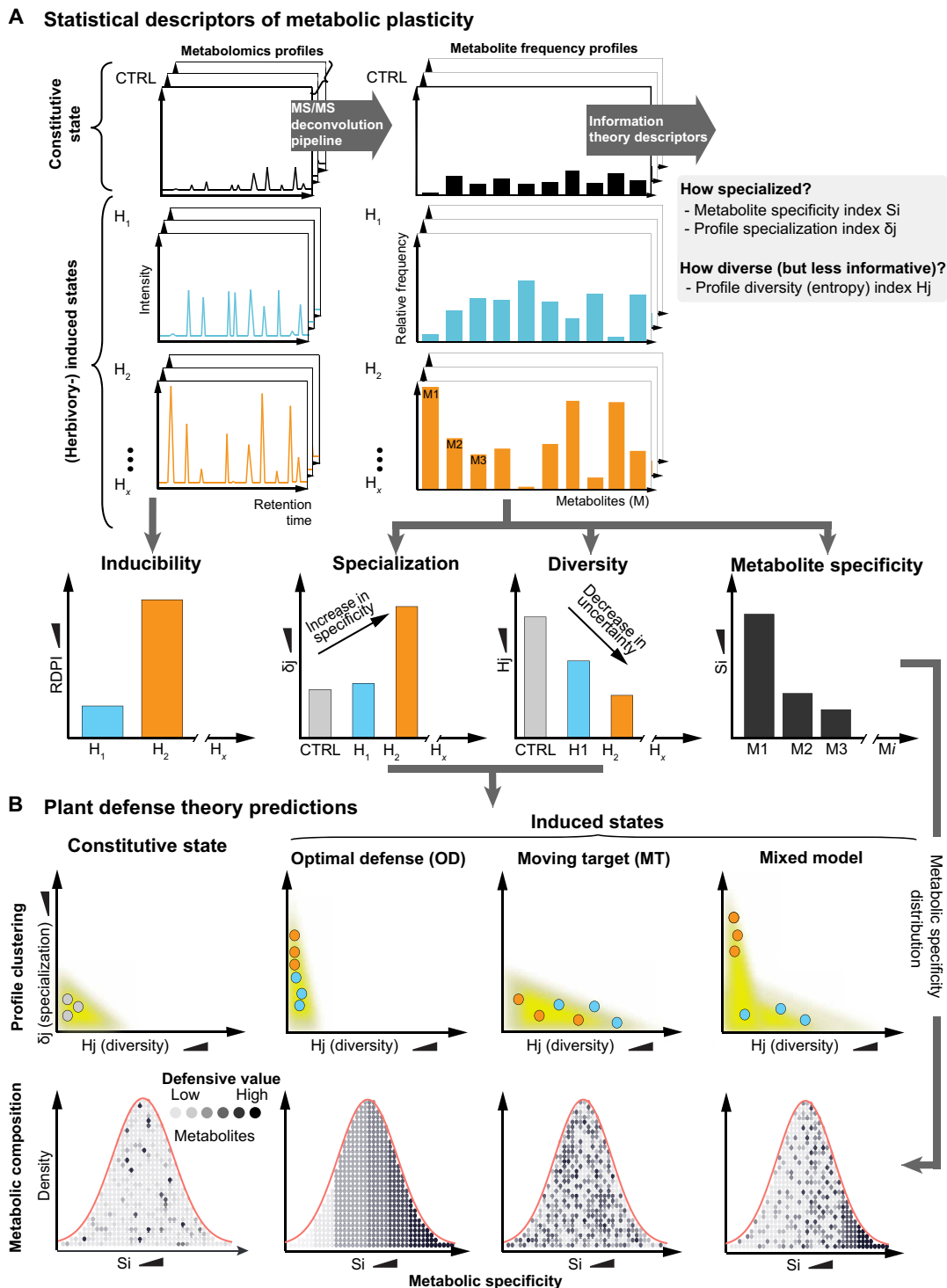


Fig. 1. Testing and contrasting the predictions of OD and MT theories using an information theory statistical framework. (A) Statistical descriptors—inducibility (RDPI), diversity (H_j index), specialization (δ_j index), and metabolite specificity (S_i index)—for herbivory-elicited (H_1 to H_x) MS/MS profiles. Increased specialization (δ_j) indicates that, on average, more herbivory-specific metabolites are produced, whereas decreased diversity (H_j) indicates either fewer produced metabolites or that metabolite distribution within profiles is less uniform. S_i value evaluates whether a metabolite is specific to a given condition (here, herbivory) or is, in contrast, maintained at the same level. (B) Conceptualized diagrams for defense theory predictions using information theory axes. OD theory predicts that herbivore attack increases defense metabolites, thereby increasing δ_j . Simultaneously, H_j decreases, as profiles are reorganized toward lowered metabolic information uncertainty. MT theory predicts that herbivore attack results in nondirectional changes in the metabolome, increasing H_j as an indicator of increased metabolic information uncertainty and creating a random distribution of S_i . We additionally propose a mixed model, the optimal MT in which some metabolites with high defensive value specifically increase (high S_i values), while others exhibit random responses (lower S_i values).

metabolome specialization (δ_j index) driven by increases in the metabolic specificity (S_i index) of certain groups of specialized metabolites with high defense value and (ii) a decrease in metabolome diversity (H_j index) caused by changes in the metabolic frequency profiles to more leptokurtic distributions. At the individual metabolite level, an orderly S_i distribution in which metabolites increase S_i values according to their defense values is expected (Fig. 1B). Along this line, we interpret the MT theory to predict that elicitation will result in (i) a decrease in δ_j index as a result of nondirectional changes in metabolites and (ii) an increase in the H_j index caused by increasing levels of metabolic uncertainty or randomness, the generalized form of diversity that can be quantified by Shannon entropy. As for the metabolic composition, the MT theory would predict a random distribution of S_i . Considering that some metabolites are specific to certain conditions but nonspecific to others and their defense value is context dependent, we additionally propose a mixed defense model, in which δ_j and H_j increase in both directions following the S_i distribution that only certain groups of metabolites, which have high defensive values, will specifically increase in S_i , while others will have a random distribution (Fig. 1B).

Information theory–analyzed reconfigurations of leaf metabolomes by insect feeding follow OD theory predictions

To test the reformulated predictions of the defense theories in the axes of information theory descriptors, we reared larvae of either the specialist (Ms) or a generalist (Sl) herbivore on leaves of rosette-stage *N. attenuata* plants (Fig. 2A). Using MS/MS analysis, we retrieved 599 nonredundant MS/MS spectra from methanolic extracts of leaf tissues collected after caterpillar feeding (data file S1). Visualizing reconfigurations of the information content in MS/MS profiles using the RDPI, H_j and δ_j indices revealed intriguing patterns (Fig. 2B). A general trend was the time-dependent increases in all degrees of metabolic reorganizations as described by the information descriptors as caterpillars continuously fed on leaves: 72 hours after herbivore feeding, RDPI was strongly enhanced; H_j was significantly decreased compared with undamaged controls as a result of an increase in specialization of the metabolic profile as quantified by the δ_j index. This clear trend was in agreement with the predictions of OD theory and inconsistent with the main predictions of MT theory, which posits that stochastic (nondirectional) changes of metabolite levels are used as a defensive camouflage (Fig. 1B). Direct feeding by the two herbivores, albeit differing in their oral secretion (OS) elicitor contents and feeding behaviors (31), resulted in similar directional changes in H_j and δ_j at both 24- and 72-hour harvests. The only discrepancy occurred in RDPI at 72 hours for which Sl feeding elicited a greater overall metabolic inducibility compared to that elicited by Ms feeding.

To explore whether these metabolome-level herbivory-induced reconfigurations are reflected in changes at the level of individual metabolites, we first focused on metabolites which were previously investigated in *N. attenuata* leaves with proven antiherbivory functions. Phenolamides are hydroxycinnamic-polyamine conjugates that accumulate during insect herbivory and are known to decrease insect performance (32). We retrieved the precursors of the corresponding MS/MSs and plotted their accumulation kinetics (fig. S1). As expected, phenolic derivatives such as chlorogenic acid (CGA) and rutin, which are not directly involved in antiherbivory defense, were down-regulated after herbivory. In contrast, phenolamides were highly

enhanced by herbivory. The continuous feeding by the two herbivores resulted in almost identical elicitation profiles of phenolamides, a pattern that was particularly apparent for the de novo synthesized phenolamides. The same phenomenon was observed when exploring the 17-hydroxygeranylinalool diterpene glycosides (17-HGL-DTGs) pathway that produces abundant acyclic diterpenes with potent antiherbivore functions (33), in which similar expression profiles were triggered by Ms and Sl feeding (fig. S1).

Simulated herbivory recapitulates directionality and species specificity of leaf metabolome changes

Possible drawbacks of direct herbivore feeding experiments are the herbivore-specific differences in leaf consumption rates and timing of feeding that make it difficult to disentangle wounding and herbivore-specific effects induced by herbivory. To better resolve the herbivore species specificity of induced leaf metabolic responses, we mimicked Ms and Sl larval feeding by immediately applying freshly collected OSs (OS_{Ms} and OS_{Sl}) to standardized puncture W in leaves at consistent leaf positions. This procedure, referred to as the W + OS treatment, standardizes the elicitation by precisely timing the initiation of herbivory elicited responses without the confounding effects of differences in rates or amounts of tissue loss (Fig. 2A) (34). Using the MS/MS analytical and computational pipeline, we retrieved 443 MS/MS spectra (data file S1), which largely overlapped with those previously assembled from the direct feeding experiment. An information theory analysis of this MS/MS dataset revealed that the reprogramming of the leaf specialized metabolome by simulated herbivory exhibited OS-specific elicitation (Fig. 2C). In particular, OS_{Ms} elicited a stronger increase in metabolome specialization at 4 hours than did the OS_{Sl} treatment. Notably, metabolic kinetics visualized in a two-dimensional space using H_j and δ_j as coordinates were consistent with a directional increase over time of metabolome specialization in response to simulated herbivory treatments as compared to the direct herbivore feeding experiment dataset (Fig. 2D). In parallel, we quantified levels of amino acids, organic acids, and sugars (data file S2) to investigate whether this directional increase in metabolome specialization in response to simulated herbivory resulted from reconfigurations of central carbon metabolism (fig. S2). To better interpret this pattern, we further monitored metabolic accumulation kinetics of the previously discussed phenolamide and 17-HGL-DTG pathways. The herbivore OS-specific induction translated into differential reconfiguration patterns within phenolamide metabolism (fig. S3). Phenolamides containing coumaroyl and caffeoyl moieties were preferentially induced by OS_{Sl} elicitation, while OS_{Ms} triggered a specific induction of feruloyl conjugates. For the 17-HGL-DTG pathway, an OS differential elicitation effect was detected for the downstream malonylated and dimalonylated products (fig. S3).

An early priming of transcriptome specialization underlies metabolome specialization

We next investigated OS-elicited transcriptome plasticity using a time-course microarray dataset that simulated herbivory using OS_{Ms} treatments to leaves of rosette-stage *N. attenuata* plants; the sampling kinetics largely overlapped with those used in the present metabolomics study (35). Compared with metabolome reconfigurations in which metabolic plasticity particularly increased over time, we observed a transient burst of transcription in leaves induced by Ms, in which transcriptome inducibility (RDPI) and specialization (δ_j)

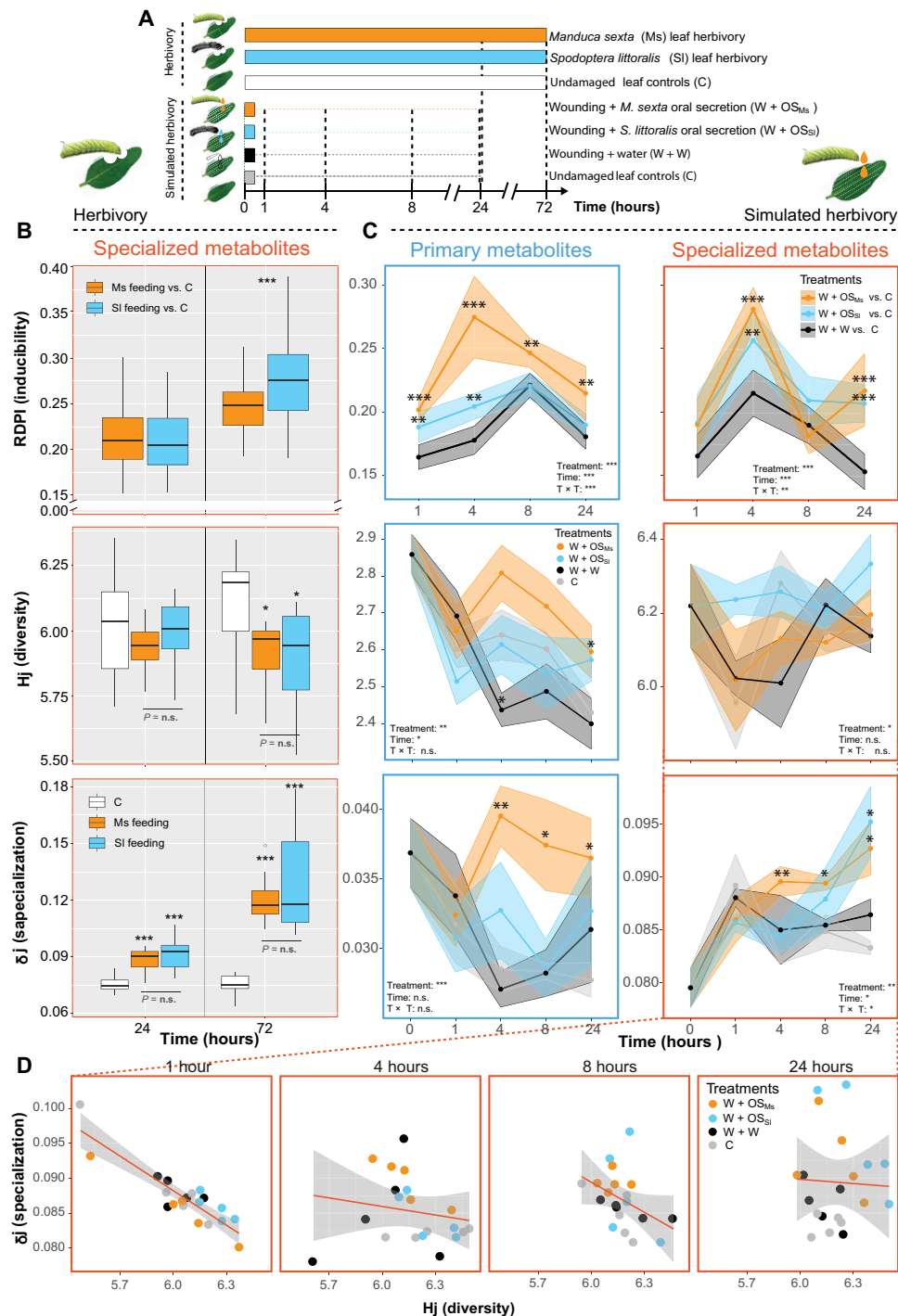


Fig. 2. Insect feeding and simulated herbivory reprograms herbivore-specific trajectories of leaf specialized and primary metabolites. (A) Experimental design: *N. attenuata* leaves were fed on by a generalist (SI) or specialist (Ms) herbivore, whereas, for simulated herbivory, puncture wounds of standardized leaf positions were treated with OSs of Ms (W + OS_{Ms}) and SI (W + OS_{SI}) larvae or water (W + W). Controls (C) are undamaged leaves. (B) Inducibility (RDPI compared to control profiles), diversity (Hj index) and specialization (δj index) indices calculated for specialized metabolite profiles (599 MS/MS; data file S1). Asterisks indicate significant differences between direct herbivore feeding treatments and controls (Student's *t* tests on pairwise differences, **P* < 0.05 and ****P* < 0.001). n.s., not significant. (C) Time-resolved indices for primary (blue box, amino acids, organic acids, and sugars; data file S2) and specialized metabolite profiles (red box, 443 MS/MS; data file S1) after simulated herbivory treatments. Ribbons refer to 95% confidence intervals. Asterisks indicate significant differences between treatments and controls [two-way analysis of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) post hoc multiple comparisons, **P* < 0.05, ***P* < 0.01, and ****P* < 0.001]. (D) Scatterplot of diversity and specialization of specialized metabolite profiles (replicated samples of different treatments).

were strongly enhanced at 1 hour, whereas diversity (Hj) was strongly decreased at this time point, followed by a relaxation of transcriptome specialization (fig. S4). Metabolic gene families such as P450, glycosyltransferases, and BAHD acyltransferases involved in the assembly of specialized metabolites from building blocks derived from primary metabolism followed the early high specialization pattern described above. The phenylpropanoid pathway was analyzed as a case study, and this analysis confirmed that core genes in phenolamide metabolism were highly OS-induced and tightly core-configured in their expression patterns during herbivory compared

to those of unelicited plants. The transcription factor, *MYB8*, and structural genes, *PAL1*, *PAL2*, *C4H*, and *4CL*, in the upstream part of this pathway exhibited early priming of transcriptions. Acyltransferases such as *AT1*, *DH29*, and *CV86* that function in the final assembly of phenolamides exhibited prolonged up-regulated patterns (fig. S4). The above observations suggest that the early priming of transcriptome specialization and the late enhancement of metabolite specialization are coupled patterns, likely as a result of synchronized regulatory systems that launch robust defense responses.

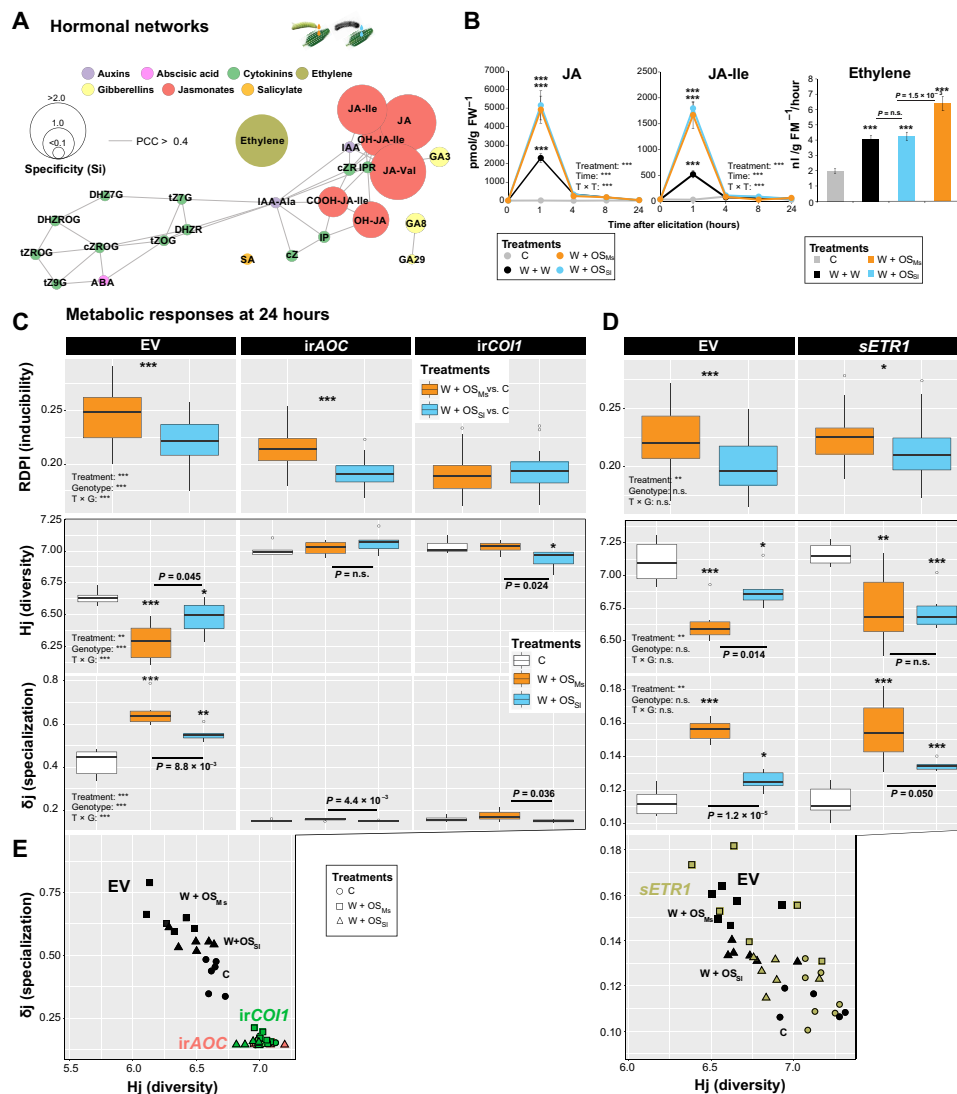


Fig. 3. Phytohormone signaling shapes the herbivore-specific metabolic information content of elicited leaves. (A) Coexpression network analysis based on PCC calculations of simulated herbivory-induced phytohormone accumulation kinetics. Nodes represent individual phytohormone, and node size indicates Si indices of phytohormone specificity among treatments. (B) JA, JA-Ile, and ET accumulation in leaves elicited by different treatments indicated with different colors: apricot, W + OS_{M_s}; blue, W + OS_{S_i}; black, W + W; gray, C (controls). Asterisks indicate significant differences between treatments and controls (two-way ANOVA followed by Tukey's HSD post hoc multiple comparisons, ****P* < 0.001). Information theory analysis of (C) 697 MS/MSs (data file S1) in JA biosynthesis and perception-impaired lines (*irAOC* and *irCOI1*) and of (D) 585 MS/MSs (data file S1) in the ET signaling-impaired *ETR1* line and in empty vector (EV) control plants elicited by the two simulated herbivory treatments. Asterisks indicate significant differences between W + OS treatments and undamaged controls (two-way ANOVA followed by Tukey's HSD post hoc multiple comparisons, **P* < 0.05, ***P* < 0.01, and ****P* < 0.001). (E) Scatterplots of diversity against specialization. Colors denote different transgenic lines; symbols denote the different treatments: triangle, W + OS_{S_i}; rectangular, W + OS_{M_s}; circle, C.

Phytohormone signaling shapes herbivore-specific changes in the information content of leaf metabolic profiles

Reconfigurations in phytohormonal signaling act as regulatory layers that integrate herbivory information to reprogram a plant's physiology. We measured accumulation kinetics for key phytohormonal classes after herbivory simulation and visualized temporal coexpression among these [Pearson correlation coefficient (PCC) of >0.4] (Fig. 3A). As expected, phytohormones that are biosynthetically related were linked within the phytohormone coexpression networks. Furthermore, metabolic specificity (Si index) was mapped onto this network to highlight phytohormones that were differentially induced by the different treatments. Two dominant regions of herbivory-specific responses were mapped: one within the JA cluster, in which JA, its bioactive form JA-Ile, and other JA derivatives exhibited the highest Si scores; and another one was for ethylene (ET). Gibberellins exhibited only moderate increases in herbivore specificity, while other phytohormones such as cytokinins, auxins, and abscisic acid exhibited low specificity to the herbivore elicitation. Strong specificity indices for JAs essentially translated from the amplification of peaking values of JA derivatives by OS application (W + OS) compared with W + W alone. Unexpectedly, OS_{M_s} and OS_{S_i}, which are known to differ in their elicitor contents (31), induced similar accumulations of JA and JA-Ile. ET was specifically and strongly induced by OS_{M_s} application, contrasting with OS_{S_i}, which did not amplify the basal wound response (Fig. 3B).

Next, we used *N. attenuata* transgenic lines modified in key steps in JA and ET biosynthesis (*irAOC* and *irACO*) and perception (*irCOII* and *sETR1*) to analyze the relative contribution of these two phytohormones to herbivory-induced metabolic reprogramming. Consistent with the previous experiments, we confirmed a general decrease in H_j index concomitant with an increase in δ_j index as a result of herbivore-OS elicitation in empty vector (EV) plants (Fig. 3, C to D) and that OS_{M_s}-elicited responses were more pronounced than those triggered by OS_{S_i}. Specific deregulations were visualized using biplots with H_j and δ_j as coordinates (Fig. 3E). The most obvious trend was that herbivory-elicited changes in metabolome diversity and specialization were almost fully erased in JA signaling-deficient lines (Fig. 3C). In contrast, silenced ET perception in *sETR1* plants, while having a much lower impact on the overall magnitude of herbivory-metabolic changes than JA signaling, attenuated differences in H_j and δ_j indices between OS_{M_s} and OS_{S_i} elicitation (Fig. 3D and fig. S5). This suggests that ET signaling serves as the fine-tuner of the herbivore species-specific metabolic responses in addition to the core function of JA signaling. Consistent with this fine-tuning function, overall metabolome inducibility was not altered in *sETR1* plants. On the other hand, *irACO* plants induced a similar overall magnitude of herbivory-elicited metabolic changes compared to those of *sETR1* plants but exhibited significantly different H_j and δ_j scores between OS_{M_s} and OS_{S_i} elicitation (fig. S5).

Exploring MS/MS herbivore-specific associations using MS/MS structural analysis

To pinpoint specialized metabolites with significant contributions to herbivore species-specific responses and whose production was fine-tuned by ET signaling, we used a previously developed structural MS/MS approach. This approach relies on a biclustering method for the de novo inference of metabolic families from MS/MS fragment [normalized dot product (NDP)] and neutral loss (NL)-based similarity scoring. The MS/MS dataset constructed from the analysis

of ET transgenic lines resulted in 585 MS/MSs (data file S1), which was resolved by biclustering into seven main MS/MS modules (M) (Fig. 4A). Some of these modules are densely populated with previously characterized specialized metabolites: for instance, M1, M2, M3, M4, and M7 were enriched with various phenolic derivatives (M1), flavonoid glycosides (M2), acyl sugars (M3 and M4), and 17-HGL-DTGs (M7). Furthermore, metabolic specificity information (Si index) was also calculated for individual metabolites in each module for which the Si distributions were visualized. Briefly, MS/MS spectra that exhibited high herbivory-elicited and genotypic specificity were characterized by high Si values with kurtosis statistics showing right-tailed leptokurtic distributions. One such leptokurtic distribution was detected for M1 within which phenolamides exhibited the highest Si scores (Fig. 4B). Previously mentioned herbivory-inducible 17-HGL-DTGs within M7 exhibited medium Si scores, indicative of a moderate differential regulation by the two OS types. In contrast, mostly constitutively produced specialized metabolites, such as rutin, CGA, and acyl sugars, were among the lowest Si scores. To better explore the structural intricacies among specialized metabolites and their Si distributions, molecular networks were constructed for each of the modules (Fig. 4B). One critical prediction of OD theory (summarized in Fig. 1B) is that the reorganization of specialized metabolites after herbivory should lead to an unidirectional change in metabolites that have high defensive values, particularly through increases in their specificity, a pattern contrasting to the random distribution of defensive metabolites predicted by the MT theory. Most phenolic derivatives clustered in M1 have been functionally associated with decreases in insect performance (32). When comparing Si values within M1 metabolites between induced and constitutive leaves of EV control plants at 24 hours, we observed a clear trend of increase in metabolic specificity for many of the metabolites after insect herbivory (Fig. 4C). Specific increases in Si values were exclusively detected for defensive phenolamides but not for other phenolics and yet unknown metabolites coexisting within this module, a specialization pattern that is in agreement with the central predictions of OD theory regarding herbivory-induced metabolic changes. To test whether this specialization of the phenolamide profile translated from OS-specific ET induction, we plotted metabolite Si indices with differential expression values elicited between OS_{M_s} and OS_{S_i} in EV and *sETR1* genotypes (Fig. 4D). In *sETR1*, divergences in phenolamide induction between OS_{M_s} and OS_{S_i} were largely attenuated. The biclustering approach was also applied to MS/MS data collected in JA-deficient lines to infer main MS/MS modules associated with the JA-regulated metabolic specialization (fig. S6).

Natural variation in JA bursts underlie intraspecific variations in herbivory-induced metabolome specialization

We further extended the analysis from a single *N. attenuata* genotype to natural populations in which intense intraspecific variations in herbivory-induced JA levels and specialized metabolites levels have been previously described (26). Using this dataset that covers 43 accessions consisting of 123 individual *N. attenuata* plants derived from seeds collected at different native habitats in Utah, Nevada, Arizona, and California (fig. S7), we calculated metabolome diversity (here referred to as population-level β diversity) and specialization shaped by OS_{M_s} elicitation. Consistent with the previous study, we observed a wide range of metabolic variations along H_j and δ_j axes, indicating that accessions differ markedly in the plasticity of their

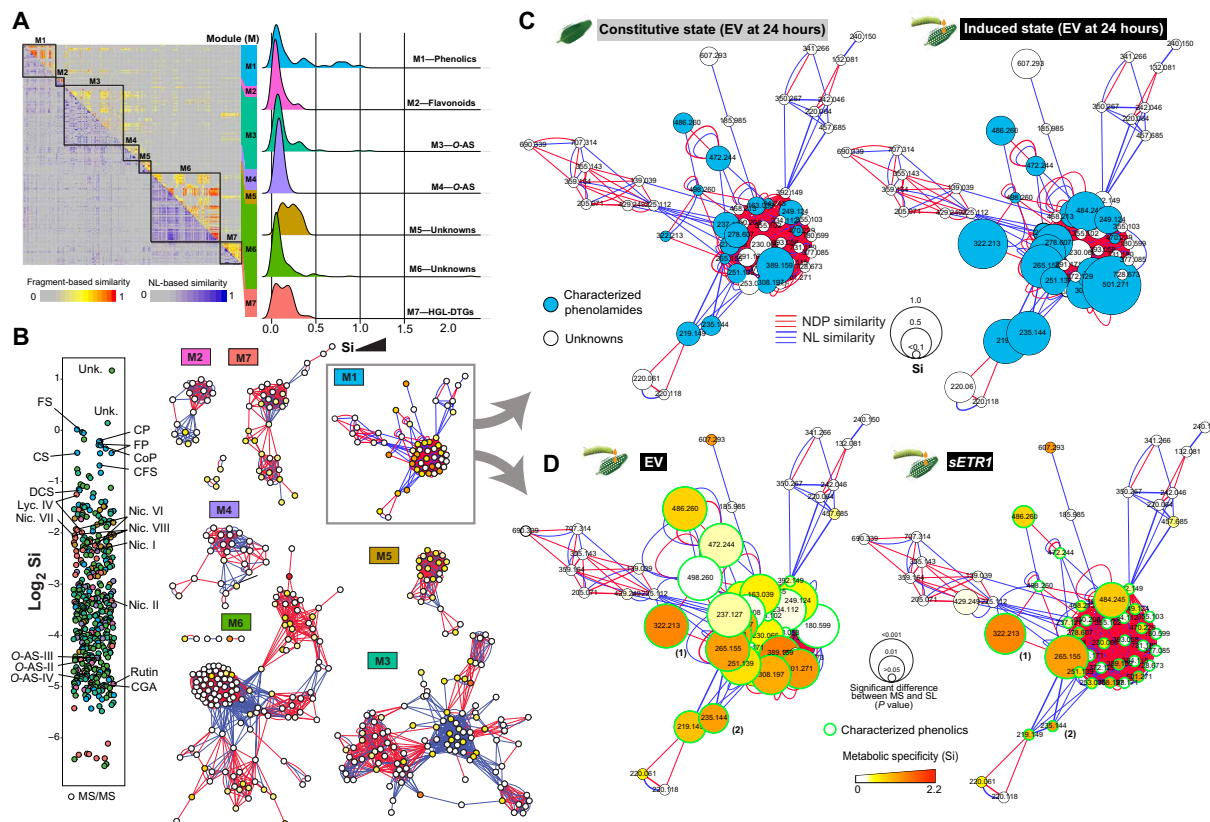


Fig. 4. MS/MS structural classification and metabolic specificity information reveals a phenolic-enriched module contributing most to the insect species-specific responses. (A) Biclustering of 585 MS/MSs based on shared fragments (NDP similarity) and shared neutral losses (NL similarity) results into modules (M) congruent with known compound families or composed of yet unknown or poorly characterized metabolites. Metabolite (MS/MS) specificity (S_i) distribution is depicted next to each module. (B) Module molecular networks: Nodes represent MS/MSs and edges, NDP (red) and NL (blue) MS/MS specificity (cutoff, >0.6). Ranked metabolite specificity indices (S_i) colored based on modules (left) and mapped onto the molecular networks (right). (C) Module M1 at constitutive (controls) and induced state (simulated herbivory) in EV plants at 24 hours: mapping of the molecular network (S_i values as node sizes and defensive phenolamides highlighted in blue). (D) Mapping of the M1 molecular network for EV and the ET perception impaired line, *sETR1*: characterized phenolics as green-circled nodes, and degree of significant differences (P value) between $W + OS_{M_s}$ and $W + OS_{S_l}$ treatments as node sizes. CP, *N*-caffeoyl-putrescine; CS, *N*-caffeoyl-spermidine; FP, *N*-feruloyl-putrescine; FS, *N*-feruloyl-spermidine; CoP, *N'*,*N''*-coumaroyl-putrescine; DCS, *N'*,*N''*-dicafeoyl-spermidine; CFS, *N'*,*N''*-cafeoyl, feruloyl-spermidine; Lyc., lyciumoside; Nic., nicotianoside; O-AS, O-acyl sugars.

metabolic responses to herbivory (fig. S7). This organization is reminiscent of previous observation made on the dynamic range of variation in herbivory-induced levels of JAs and that extreme values are maintained within single populations (26, 36). By testing population-level correlations between H_j and δ_j with JA and JA-Ile, we found significant positive correlations between JAs and both metabolome β diversity and specialization indices (fig. S7). This suggested that the herbivory-induced heterogeneity in the elicitation of JAs detected at the population level is likely a key metabolic polymorphism resulting from selection from insect herbivores.

Exploring assignments of OS-elicited responses in closely related *Nicotiana* species reveals species-specific coordination of specialized metabolites with JA plasticity

Previous research has shown that *Nicotiana* species largely differ in the type and relative reliance on inducible versus constitutive metabolic defenses. These variations in antiherbivore signaling and defenses are thought to be modulated by insect population pressures, plant life cycles, and the costs of defense production within the ecological niche in which a given species grows. We examined the

consistency of herbivory-induced reconfigurations of leaf metabolomes in six *Nicotiana* species, native to North and South America, and closely related to *N. attenuata*, namely, *Nicotiana pauciflora*, *Nicotiana miersii*, *N. attenuata*, *Nicotiana acuminata*, *Nicotiana linearis*, *Nicotiana spegazzinii*, and *Nicotiana obtusifolia* (Fig. 5A) (37). Six of these species, including the well-characterized species *N. attenuata*, are annuals from the *Petunioides* clade, while *N. obtusifolia* is a perennial plant from the sister clade *Trigonophyllae* (38). These seven species were subsequently subjected to $W + W$, $W + OS_{M_s}$, and $W + OS_{S_l}$ induction to study species-level metabolic reconfigurations to insect herbivory.

Using the biclustering approach, we identified nine modules of 939 MS/MSs (data file S1). The MS/MS compositions that were reconfigured by different treatments differed greatly in different modules among species (fig. S8). Visualizing H_j (here referred to as species-level γ diversity) and δ_j revealed that the different species clustered as very distinct groups in the metabolic space, with the species-level demarcation effect being frequently more prominent than the elicitation effects. Exceptions were *N. linearis* and *N. spegazzinii*, which exhibited broad dynamic ranges of elicitation effects (Fig. 5B).

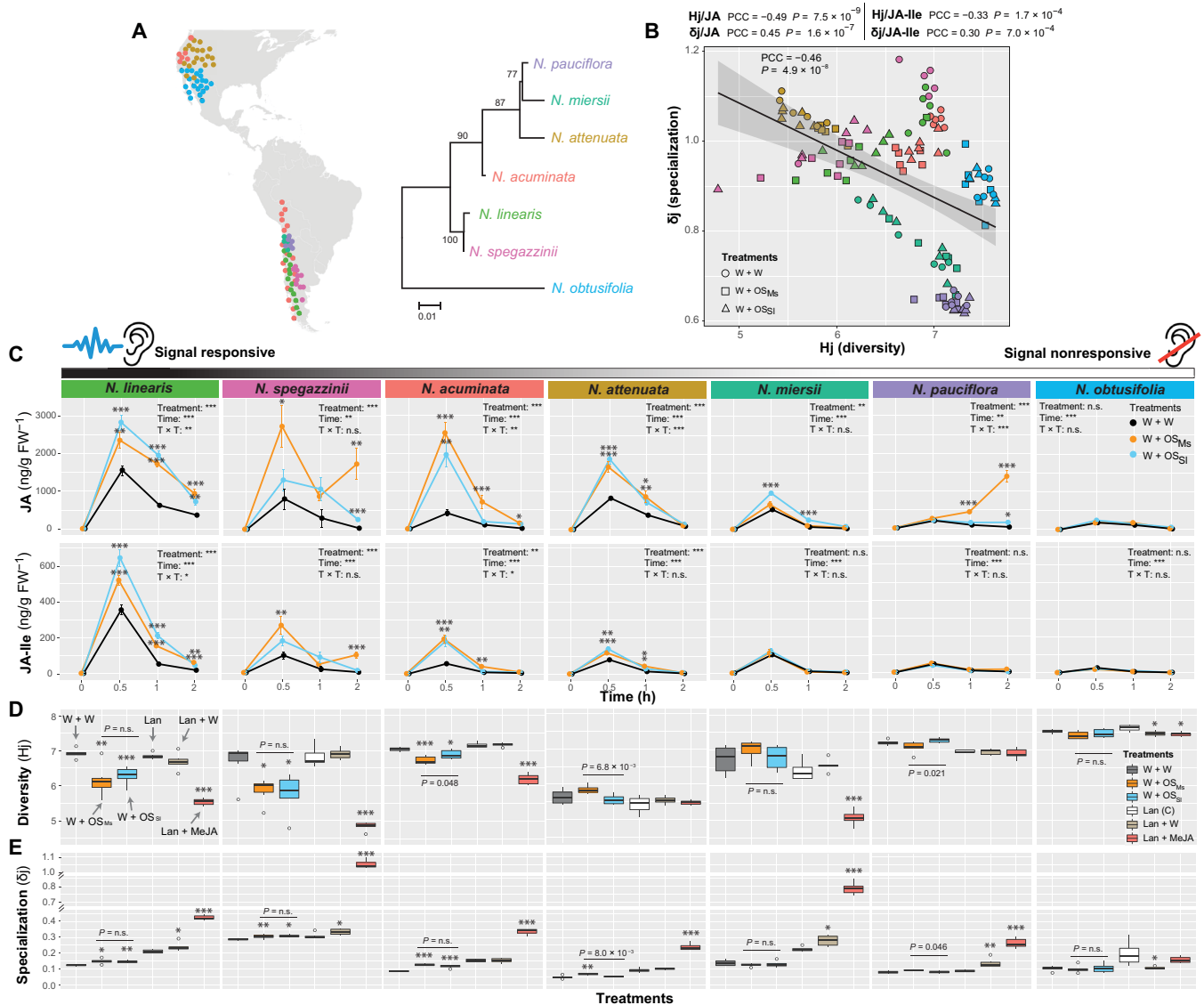


Fig. 5. OS-elicited specialized metabolite responses in closely related *Nicotiana* species reveal species-specific trajectories correlated with plasticity in JA signaling. (A) Maximum likelihood-based bootstrapped phylogenetic tree [for nuclear glutamine synthetases (38)] and geographic distribution (37) of seven closely related *Nicotiana* species (different colors). (B) Scatterplots of diversity against specialization for the metabolic profiles of the seven *Nicotiana* species (939 MS/MS; data file S1). At the species level, metabolome diversity negatively correlates with specialization. Species-level correlation analyses of metabolome diversity and specialization with JA accumulation are shown in fig. S9. Colors, different species; triangles, W + OS_{Sj}; rectangles, W + OS_{Ms}; circles, W + W. (C) *Nicotiana* species JA and JA-Ile kinetics ranked according to OS elicitation amplitude (two-way ANOVA followed by Tukey's HSD post hoc multiple comparisons, * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$ for the W + OS and W + W comparison). Box plots of (D) diversity and (E) specialization for each species after simulated herbivory and methyl JA (MeJA). Asterisks indicate significant differences between W + OS and W + W or between lanolin plus W (Lan + W) or Lan plus MeJA (Lan + MeJA) and Lan controls (two-way ANOVA followed by Tukey's HSD post hoc multiple comparisons, * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$).

In contrast, species such as *N. pauciflora* and *N. obtusifolia* exhibited less pronounced metabolic responses to treatments but greater metabolome diversity. The species-specific distributions of induced metabolic responses resulted in a significant negative correlation between specialization and γ diversity (PCC = -0.46, $P = 4.9 \times 10^{-8}$). Variations in OS-induced levels of JAs were positively correlated with metabolome specialization but negatively correlated with metabolome γ diversity exhibited by each species (Fig. 5B and fig. S9). Notably, species colloquially referred to as “signal responsive” ones in Fig. 5C, such as *N. linearis*, *N. spegazzinii*, *N. acuminata*, and *N. attenuata*

elicited remarkable OS-specific JA and JA-Ile burst at 30 min, while others referred to as “signal nonresponsive” ones such as *N. miersii*, *N. pauciflora*, and *N. obtusifolia* showed only marginal inductions of JA-Ile and without any OS specificity (Fig. 5C). At the metabolic level and as previously described for *N. attenuata*, signal responsive species exhibited an OS-specific and significant increase δ_j concomitant with a decrease Hj. This OS-specific elicitation effect was not detected in species classified as signal nonresponsive ones (Fig. 5, D and E). The OS-specifically elicited metabolites are shared more frequently among signal responsive species, which clustered apart

from the signal nonresponsive ones that exhibited less species interdependencies (fig. S8). This result suggested that the OS-specific induction of JAs and the OS-specific reconfigurations of downstream metabolomes were coupled patterns at the species level.

We next investigated whether these coupling patterns were constrained by JA availability using exogenous JA application by treating plants with lanolin paste containing methyl JA (MeJA), which is rapidly de-esterified to JA in the plant's cytoplasm. We found the same trend of graded changes from signal-responsive to signal-nonresponsive species induced by the continuous supply of JA (Fig. 5, D and E). Briefly, the MeJA treatment strongly reprogrammed the metabolomes of *N. linearis*, *N. spegazzinii*, *N. acuminata*, *N. attenuata*, and *N. miersii* toward pronounced increases in δj and decreases in H_j . *N. pauciflora* only showed an increase in δj but not H_j . *N. obtusifolia*, which was previously shown to accumulate extremely low levels of JAs, was also poorly responsive to MeJA treatment in terms of metabolome reconfigurations. These results suggested that JA production or signal transduction is physiologically constrained in the signal nonresponsive species. To test this hypothesis, we investigated the transcriptomes of four of the species (*N. attenuata*, *N. miersii*, *N. pauciflora*, and *N. obtusifolia*) induced by W + W, W + OS_{Me}, and W + OS_{SI} (39). Consistent with the pattern of metabolome reconfigurations, species were well separated in the transcriptomic space with *N. attenuata* exhibited the highest OS-induced

RDPI while *N. obtusifolia* being the lowest (Fig. 6A). However, the transcriptome diversity induced by *N. obtusifolia* was found to be the lowest among the four species, a pattern opposite to the highest metabolome diversity of *N. obtusifolia* previously shown among seven species. Previous research has revealed that a group of genes involved in early defense signaling, including JA signaling, accounted for the specificity of herbivore-associated elicitors induced early defense responses within *Nicotiana* species (39). Comparing the JA signaling pathway among the four species revealed interesting patterns (Fig. 6B). A majority of genes in this pathway, such as AOC, OPR3, ACX, and CO11, exhibited comparably high induction levels among the four species. However, a key gene, JAR4, which converts JA to its bioactive form JA-Ile accumulated transcripts at very low levels specifically in *N. miersii*, *N. pauciflora*, and *N. obtusifolia*. In addition, transcripts of another gene, AOS, were not detected only in *N. obtusifolia*. These changes in gene expression are likely to be responsible for the low JA production in the signal nonresponsive species and the extreme phenotype of constrained induction for *N. obtusifolia*.

Impact of species-specific induced responses on herbivore resistance

In this last section, we examine how the insect species-specific remodeling of metabolomes of different plant species contributed

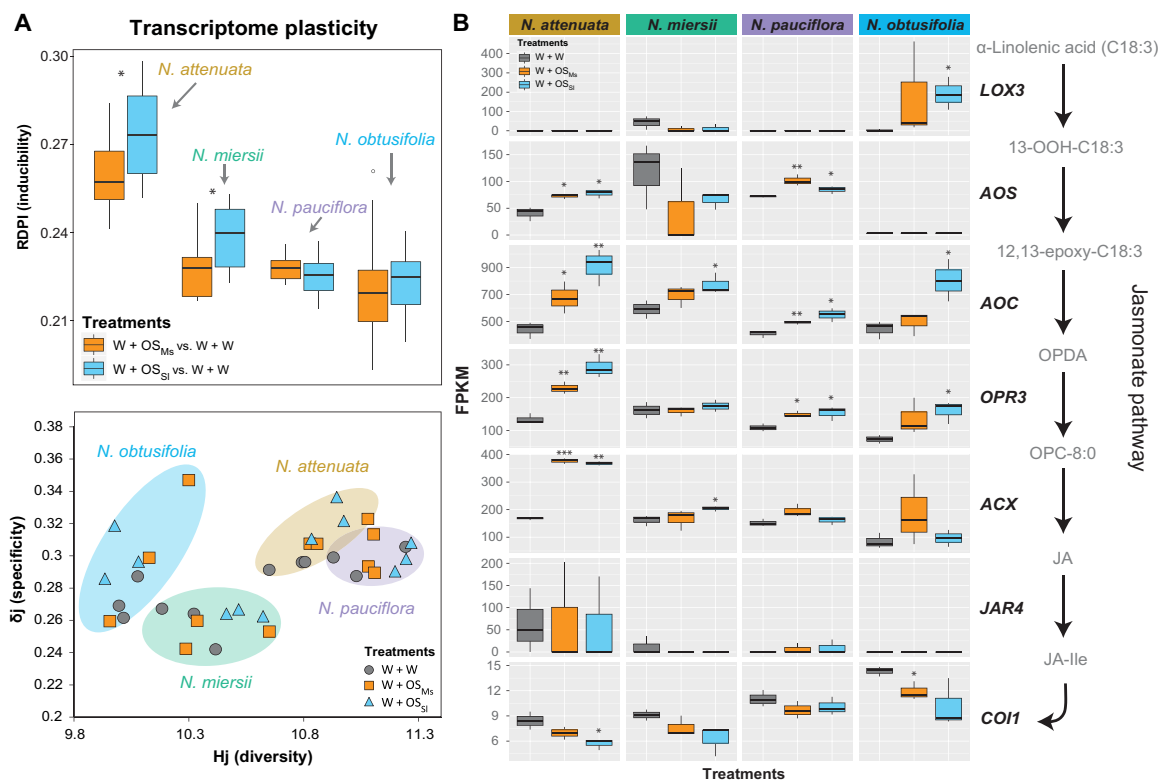


Fig. 6. OS-elicited transcriptome responses in closely related *Nicotiana* species highlight key genes in the JA signaling pathway responsible for the plant species-specific responses to herbivory. (A) Information theory analysis of the reprogramming of early transcriptomic responses of four closely related *Nicotiana* species sampled 30 min after herbivory induction. RDPI was calculated by comparing herbivore OS-elicited leaves with wounding controls. Colors indicate different species and symbols indicate different treatments. (B) Gene expression analysis in the JA signaling pathway among the four species. A simplified JA pathway is shown next to box plots. Different colors indicate different treatments. Asterisks indicate significant differences between W + OS treatments and W + W controls (Student's *t* tests on pairwise differences, **P* < 0.05, ***P* < 0.01, and ****P* < 0.001). OPDA, 12-oxo-phytodienoic acid; OPC-8:0, 3-oxo-2(2'(Z)-pentenyl)-cyclopentane-1-octanoic acid.

to herbivore resistance. Previous research has highlighted that *Nicotiana* spp. differed greatly in their induced resistance to Ms and Sl larval performance (40). Here, we investigated how this pattern was linked to their metabolic plasticity. Using the above-mentioned four *Nicotiana* species and testing the correlation between herbivory-induced metabolome diversity and specialization and the plants' resistance to Ms and Sl, we found that the resistance to the generalist Sl was positively correlated with both diversity and specialization, whereas the resistance to the specialist Ms showed weaker correlations with specialization and was not significantly correlated with diversity (fig. S10). For resistance to Sl, both *N. attenuata* and *N. obtusifolia*, previously shown to differ greatly in their responses to herbivore induction at both the level of JA signaling and metabolome plasticity, showed similarly high levels of resistance.

DISCUSSION

Over the past six decades, plant defense theories have provided theoretical frameworks from which researchers have made predictions about the evolution and function of the considerable diversity of the specialized metabolites of plants. Most of these theories have not followed the normal procedures of strong inference (41), presenting critical predictions posed at the same levels of analysis (3), which would allow the field to move forward when the tests of critical predictions allow a particular theory to be supported, while rejecting others (42). Instead, new theories made predictions at different levels of analysis, adding new descriptive layers of consideration (42). However, two theories, MT and OD, posed at the functional level of analysis, could be readily interpreted as making critical predictions regarding herbivore-induced changes in specialized metabolism: OD theory predicted that the changes in specialized metabolism "space" would be highly directional, favoring metabolites with high defensive value, while the MT theory posited that the changes would be nondirectional and randomly positioned in the metabolic space. Previous examinations of the predictions of the OD and MT have been tested with a narrow set of a priori "defense" compounds. These metabolite-focused tests preclude the analysis of the extent and trajectory of metabolome reconfigurations during herbivory and do not allow for a consistent statistical framework within which to test these critical predictions that require the ability to quantify changes in a plant's metabolome in holistic terms. Here, we used innovations in computational MS-based metabolomics and deconvoluted MS analyses in the common currencies of information theory descriptors to test the critical predictions posed at the global metabolome level that distinguish the two theories. Information theory has been applied in many fields, particularly in ecology in the context of biodiversity and trophic flow studies (43). However, as far as we know, this is the first application to describe plants' metabolic information space and address ecological questions regarding the function of temporal metabolic changes in response to environmental cues. In particular, the power of the present approach lies in its ability to compare patterns within and among plant species to examine how herbivory has shaped plant metabolism at different scales in the evolutionary hierarchy, from microevolutionary patterns within species to among-species macroevolutionary patterns.

Principal components analysis (PCA) that transforms multivariable dataset into a reduced dimensionality space such that the main trends in the data are interpretable is frequently adopted as an exploratory technique to parse datasets such as deconvoluted

metabolomes. However, dimensionality reduction loses part of the information content in the dataset, and PCA provides no quantitative information about traits that are particularly germane for ecological theory, such as: How does herbivory reconfigure diversity (e.g., richness, distribution, and abundance) in specialized metabolites? Which metabolites are predictors for a given herbivory-induced state? Decomposing the information content of leaf specialized metabolite profiles in terms of specificity, diversity, and inducibility revealed that herbivore feeding activated idiosyncratic metabolic rearrangements. Unexpectedly, we observed that the resulting metabolic landscapes, as described by the implemented information theory indices, largely overlapped after attack by the two herbivores, Sl, a nocturnal feeding generalist, and Ms, a Solanaceae specialist, despite their distinct feeding behaviors and concentrations of fatty acid-amino acid conjugate (FAC) elicitors in their OS (31). Simulated herbivory treatments, by treating standardized puncture wounds with herbivore OS, revealed similar trends. This standardized procedure of mimicking a plant's responses to herbivore attack removes confounding factors caused by variations in herbivore feeding behaviors that result in different amounts of damage occurring at different times (34). FACs that are known to be the major elicitors in OS_{Ms} that induce JA and other phytohormonal responses in *N. attenuata* are hundreds of times lower in OS_{Sl} (31). However, OS_{Sl} elicited similar levels of JA accumulations compared to OS_{Ms}. The JA response in *N. attenuata* was previously shown to be very sensitive to OS_{Ms} in which FACs can retain its activity even when diluted 1:1000 with water (44). It is thus likely that the FACs in OS_{Sl}, albeit low, were sufficient to elicit a full induction of JA burst compared to OS_{Ms}. Previous studies have shown that porin-like proteins (45) and oligosaccharides (46) can serve as molecular cues in OS_{Sl} that trigger plant defense responses. However, it is still unclear whether these elicitors in OS_{Sl} were responsible for the JA accumulations observed in the current study.

While few studies have described differential metabolic fingerprints induced by different herbivores or exogenous JA or SA (salicylic acid) applications (47), none have conducted a systemic exploration of herbivore species-specific perturbations in a plant's phytohormonal network and its holistic consequences for specialized metabolic profiles. The present analysis further confirmed that intrinsic hormonal network connections with other phytohormones beyond JAs shape the specificity of herbivory-induced metabolic reconfigurations. In particular, we detected a markedly larger elicitation of ET by OS_{Ms} than by OS_{Sl}. This pattern is consistent with the greater FACs contents in OS_{Ms}, which are necessary and sufficient to elicit ET bursts (48). ET's signaling functions for plant specialized metabolite dynamics in the context of plant-herbivore interactions remains sporadic and targeted to single compound groups. Moreover, the vast majority of research has used exogenous applications of ET or its precursors or various inhibitors to study ET's regulation in which these exogenous chemical applications can have numerous nonspecific side effects. To our knowledge, this study represents the first to conduct a large-scale systematic examination of ET's role in orchestrating a plant's metabolome dynamics using transgenic plants impaired in ET production and perception. The herbivore-specific induction of ET was shown to ultimately modulate the metabolome response, most notably for the herbivore-specific de novo accumulations of phenolamides as revealed by transgenic manipulations of ET biosynthetic (*ACO*) and perception (*ETR1*) genes. ET has previously been shown to fine-tune the JA-induced nicotine accumulation by regulating putrescine *N*-methyltransferases (49).

However, it is still unclear mechanistically how ET fine-tunes phenolamide induction. In addition to ET's signaling functions, investments in polyamine-based phenolamides could be modulated by competitive shunting of metabolic flux into S-adenosyl-L-methionine, the common intermediate for both ET and polyamine biosynthetic pathways. The mechanisms responsible for the modulation of phenolamide levels by ET signaling remain to be investigated in future work.

Strong focus toward specific metabolic classes due to the sheer number of specialized metabolites of unknown structures has long precluded rigorous assessments of the temporal shifts in metabolic diversity following biotic interactions. A central result from the present information theory analysis–based unbiased metabolite-derived MS/MS spectra acquisition was that herbivore feeding or simulated herbivory consistently lowers the overall metabolic diversity of leaf metabolomes while increasing their degree of specialization. This temporal increase in the specialization of metabolomes elicited by herbivory is revealed to be coupled with coordinated increases in transcriptome specialization. Features contributing most (with high S_i values) to this greater metabolome specialization were specialized metabolites with previously characterized antiherbivory functions. This pattern is consistent with the predictions of the OD theory but not with MT predictions regarding the stochasticity of the metabolome reprogramming. However, the data were also consistent with predictions of a mixed model (optimal MT; Fig. 1B), as other uncharacterized metabolites having yet unknown defensive functions may still follow a random S_i distribution.

One notable pattern further captured by this study was that different levels of evolutionary organizations, from microevolutionary levels (single plants and *Nicotiana* populations) to larger evolutionary scales (closely related *Nicotiana* species), differed markedly in their capacities to “optimally defend” against herbivores. Moore *et al.* (20) and Kessler and Kalske (1) have independently proposed to transpose the three functional levels of biodiversity originally distinguished by Whittaker (50) to constitutive and induced temporal changes in chemodiversity; these authors did not outline procedures either for large-scale metabolome data acquisition or for how metabolic diversity could be calculated from these data. In the present study, minor adjustments to Whittaker's functional categorization would consider α -metabolic diversity as the diversity of MS/MS spectra in a given plant and β -metabolic diversity as the fundamental intraspecific metabolic space for a set of populations, while γ -metabolic diversity would be the extension of the analysis to congeneric species.

JA signaling is central for a broad spectrum of metabolic responses to herbivory. However, rigorous quantitative tests on the contributions of intraspecific modulations of JA biosynthesis to metabolome diversity are lacking, and whether JA signaling functions as a general locus for stress-elicited metabolic diversification at higher macroevolutionary scales has remained elusive. We observed that herbivory-induced metabolome specialization in *N. attenuata* and variations of metabolome specialization within *N. attenuata* populations and among closely related *Nicotiana* species were systematically positively correlated with JA signaling. Furthermore, the herbivory-induced metabolic specialization on a single genotype basis was largely abolished when JA signaling was impaired (Fig. 3, C and E). As changes in the metabolic profiles of natural *N. attenuata* populations were mostly quantitative, variation in metabolic β diversity and specialization in this analysis were likely largely driven by the strong elicitation of metabolite-rich compound classes that dominating the

metabolome profile in accessions and resulting in positive correlations with JA signaling.

More revealing was the analysis of closely related *Nicotiana* species as those differed greatly in their biochemical machineries, resulting in qualitatively specialized metabolite profiles. The information theory processing of the captured metabolic profiles revealed a trade-off, exacerbated by the herbivory induction, between metabolic γ diversity and specialization. JA signaling plays a central role in this trade-off; increases in metabolome specialization, which are consistent with the main OD prediction, were positively correlated with JA signaling, while JA signaling was negatively correlated with metabolic γ diversity. These patterns suggested that a plant's OD capacities are largely defined by JA plasticity at both microevolutionary and larger evolutionary scales. Exogenous JA application experiments, to bypass JA biosynthetic deficiencies, further revealed that closely related *Nicotiana* species could be distinguished as signal responsive and signal nonresponsive species just as they were by their patterns of JA and metabolome plasticity in response to herbivory induction. Signal nonresponsive species were physiologically constrained by their inability to produce and respond to endogenous JA, likely due to mutations in a few key genes in the JA signaling pathway (AOS and *JAR4* in *N. obtusifolia*) of which transcripts were absent. This result highlights that these interspecific macroevolutionary patterns may largely be driven by variations in internal hormone perception and responsiveness.

Beyond plant-herbivore interactions, exploring metabolic diversity is relevant to all important theoretical advances in the study of organisms' adaptations to their environments and complex phenotypic trait evolution. With the increase in data volumes acquired by modern MS instruments, hypothesis testing regarding metabolic diversity can now transcend differences in individual/classes of metabolites and proceed to global analyses revealing unsuspected patterns. In this process of larger-scale analyses, an important metaphor is the idea of constructing a meaningful map from which data can be explored. Hence, an important outcome of the present combination of unbiased MS/MS metabolomics and information theory is that it provides a simple metric with which to construct a map that allows for the browsing of metabolic diversity at the different taxonomical scales, an essential requirement for the study of micro/macroevolution and community ecology.

At the macroevolutionary level, core to Ehrlich and Raven's (51) plant-insect coevolution theory is the prediction that interspecific variations in metabolic diversity are responsible for lineage diversification in plants. However, in the five decades since this seminal work was published, few tests of this hypothesis have been conducted (52), in large part due to the phylogenetic rarity of comparable metabolic characters across distant plant lineages that can be used to anchor targeted analytical measurements. The present information theory–processed MS/MS workflows enable such a taxonomic scale comparison of these macroevolutionary patterns in specialized metabolism by quantifying MS/MS structural similarities of unknown metabolites, without a priori metabolite selection, and translate these MS/MSs into a set of simple statistical indices. The process is analogous to a phylogenetic analysis, which quantifies the rate of diversification or character evolution using sequence alignment without a priori predictions.

At the biochemical level, the screening hypothesis by Firn and Jones (53) implies that metabolic diversity is maintained at different hierarchical scales to provide the raw material to exapt bioactivities

of previously extraneous or alternatively adapted metabolites. Information theory approaches provide a framework in which to quantify these evolutionary transitions in metabolite specificity that should occur during metabolite exaptation as part of the proposed evolutionary screening processes: from low to high specificity for exapted metabolites whose bioactivity becomes adaptive for a given environmental context.

In conclusion, during the early days of molecular biology when the seminal plant defense theories were developed, deductive hypothesis-driven methods were widely considered to be the only means of making scientific advances, in large part due to the technical constraints of measuring entire metabolomes. While hypothesis-driven approaches are particularly useful in choosing among alternative causal mechanisms, their ability to advance our understanding of biochemical networks is more limited compared to the computational methods currently available in contemporary data-intensive science. Consequently, theories that made predictions far beyond the reach of the available data could not be fully falsified, thereby abrogating the hypothesis formulation/testing cycle through which a research field makes progress (4). We foresee that the metabolomics computational workflow presented here could reinvigorate interest in both proximate (how) and ultimate (why) questions regarding metabolic diversity and contribute to a new era of theory-guided data science that revisits the important theories that inspired previous generations.

MATERIALS AND METHODS

Plant treatment and sample preparation

Direct herbivore feeding was conducted by rearing one second instar Ms or Sl larvae on one *N. attenuata* leaf of individual rosette-stage plants, each with 10-plant replicates. Insect larvae were clip-caged, remaining leaf tissues were collected and flash-frozen at 24 and 72 hours after infestation, and metabolites were extracted.

Simulated herbivory treatment was conducted in a highly synchronized fashion by producing, with a fabric pattern wheel, three rows of punctures onto each side of the midvein of three fully expanded leaves from plants in the rosette stage of growth and immediately applying 1:5 diluted Ms or Sl OSs to the puncture wounds with a gloved finger. One treated leaf was harvested and processed as described above. Primary metabolites and phytohormones were extracted using previously described methods (54).

For exogenous JA applications, three petioles of leaves from each of six rosette-stage plants per species were treated with 20 μ l of lanolin paste containing 150- μ g MeJA (Lan + MeJA), with 20 μ l of lanolin plus wounding treatment (Lan + W), or with 20 μ l of pure lanolin as control. Leaves were harvested at 72 hours after treatment, flash-frozen in liquid nitrogen, and stored at -80°C until use.

Four JA and ET transgenic lines, namely, *irAOC* (36), *irCOII* (55), *irACO*, and *sETR1* (48), have been previously characterized in our group. *irAOC* strongly exhibits decreased JA and JA-Ile levels, whereas *irCOII* is insensitive to JAs and exhibits increases in JA-Ile accumulations compared to EV. Similarly, *irACO* attenuates ET production, whereas ET-insensitive *sETR1* up-regulates ET production in comparison to EV.

ET measurement

ET measurements were conducted noninvasively with a photoacoustic laser spectrometer (Sensor Sense ETD-300 real-time ET sensor). Half

leaves were excised immediately after treatment and transferred to 4-ml sealed glass vials, and the headspace was allowed to accumulate over a 5-hour time period. During measurements, each vial was flushed with a flow of purified air at 2 liters/hour for 8 min, which had previously passed through a catalyzer provided by Sensor Sense to remove CO_2 and water.

Microarray and RNA sequencing data analysis

Microarray data were originally published in (35) and deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus database (accession no. GSE30287). Data corresponding to leaves elicited by W + OS_{Ms} treatments and undamaged controls were extracted for the present study. Raw intensities were \log_2 , and baseline was transformed and normalized to their 75th percentile using the R software package, before statistical analysis.

The raw RNA sequencing (RNA-seq) data of *Nicotiana* species were retrieved from the NCBI short reads archive (SRA) under the project number PRJNA301787, which was reported by Zhou *et al.* (39) and processed as described in (56). Raw data corresponding to W + W, W + OS_{Ms}, and W + OS_{Sl} treatments of *Nicotiana* species were selected for the present study analysis and processed as follows: raw RNA-seq reads were first converted to FASTQ format. HISAT2 converted FASTQ to SAM, and SAMtools converted SAM files to sorted BAM files. StringTie was used to calculate gene expression as fragments per kilobase of transcript per million reads sequenced.

Ultrahigh-performance liquid chromatography–electrospray ionization/quadrupole and time-of-flight–MS conditions for profile mode analysis and MS/MS data acquisition

An Acclaim column (150 mm by 2.1 mm; particle size 2.2 μm) with a 4-mm by 4-mm guard column of the same material was used for the analysis. The following binary gradient was used with a Dionex UltiMate 3000 ultrahigh-performance liquid chromatography (UHPLC) system: 0 to 0.5 min, isocratic 90% A [deionized water, 0.1% (v/v) acetonitrile and 0.05% formic acid], 10% B (acetonitrile and 0.05% formic acid); 0.5 to 23.5 min, gradient phase to 10% A and 90% B; 23.5 to 25 min, isocratic 10% A and 90% B. Flow rate was 400 $\mu\text{l}/\text{min}$. For all MS analyses, the column eluent was infused into an Impact II (Bruker Daltonics, Bremen, Germany) equipped with quadrupole and time-of-flight (qTOF) analyzers and fitted with an electrospray source operated in positive ionization mode (capillary voltage, 4500 V; capillary exit, 130 V; dry temperature, 200°C ; dry gas flow, 10 liters/min).

Data-independent or indiscriminant MS/MS fragmentation analysis (hereafter referred to as MS/MS) was conducted to gain structural information about the overall detectable metabolic profile. The concept of the indiscriminant MS/MS approach relies on the fact that the quadrupole is operated with a very large mass isolation window [so that quasi all mass/charge ratio (m/z) signals are considered for fragmentation]. For this, several independent analyses are performed with increasing collision-induced dissociation collision energy (CE) values since the Impact II instrument cannot create CE ramping. Briefly, samples were first analyzed by UHPLC–electrospray ionization/qTOF-MS using the single MS mode (low fragmentation condition derived from in-source fragmentation) by scanning from m/z 50 to 1500 at a repetition rate of 5 Hz. MS/MS analyses were conducted using nitrogen as collision gas and involved independent measurements at the following four different collision-induced dissociation

voltages: 20, 30, 40, and 50 eV. The quadrupole was operated throughout the measurement with the largest mass isolation window, from m/z 50 to 1500. This mass range was automatically activated by the operating software of the instrument when the precursor m/z and the isolation width are experimentally set to 200 and 0 Da, respectively. Mass fragments were scanned as in the single MS mode. Mass calibration was performed using sodium formate (50 ml of isopropanol, 200 μ l of formic acid, and 1 ml of 1 M NaOH in water). Data files were calibrated post-run on the average spectrum from a given time segment, using the Bruker high-precision calibration algorithm. Raw data files were converted to NetCDF format using the export function of the Data Analysis v4.0 software (Bruker Daltonics, Bremen, Germany). The MS/MS dataset has been deposited in the open metabolomics database MetaboLights (www.ebi.ac.uk) under accession no. MTBLS1471.

Deconvolution of compound-specific MS/MS

MS/MS assembly was achieved via correlational analysis between MS1 and MS/MS mass signals for low and high collision energies and newly implemented rules. The correlation analysis for precursor-to-product assignment for was implemented using an R script and rules were implemented using a C# script (<https://github.com/MPI-DL/indiscriminant-MS-MS-assembly-pipeline>).

To reduce false-positive errors resulting from spurious correlations from background noise due to the fact that some m/z features are only detected in a few samples, we applied “fill peaks” function of R package XCMS (use for background noise correction) to replace “NA” (not detected peak) intensities. When the fill peaks function is used, there still were many “0” intensity values in the dataset that affect the calculation of correlations; we then compared data processing results obtained with and without the fill peaks function and calculated a background noise value from the average correction estimates, and these 0 intensity values were then replaced with the calculated background value. We also only considered features with intensities that were more than three times the background value and considered these as “true peaks.” Only m/z signals with at least eight true peaks for the samples precursors (MS1) and fragments datasets were considered for PCC calculation.

A precursor mass feature is further defined if its intensity across sample significantly correlate with the decreased intensity of the same mass feature subjected to low or high collision energies and that this feature is not annotated as an isotope peak by CAMERA. The correlation analysis was then conducted by calculating all possible precursor-to-product pairs within 3 s—estimated retention time window for peak deviation. m/z values were only considered as fragments if they were lower than that of the precursor and MS/MS fragmentation occurred in the same sample position within the dataset as the precursor from which it is derived.

On the basis of these two simple rules, we excluded assigned fragments at m/z values larger than that of the identified precursor as well as based the sample position for occurrence precursor and assigned fragments. Many in-source fragmentation-generated mass features produced in the MS1 mode can also be selected as candidate precursors resulting in redundant compound MS/MS. To reduce this data redundancy, we merged spectra if their NDP similarity exceeded 0.6 and they belong to the chromatographic “pcgroup” annotated by CAMERA. Last, we merged all four CE results for precursor-to-fragment associations into a final deconvoluted composite spectrum by choosing the highest intensity peak among all candidate peaks of

the same m/z value at the different collision energies. This latter processing step is based on the composite spectrum concept and accounts for the different CE conditions required to maximize fragmentation possibilities since certain fragments are detected only at specific collision energies.

Information theory framework for defining metabolome diversity and specialization and metabolic specificity

Metabolic profile inducibility was calculated using RDPI (30). Metabolic profile diversity, the H_j index, was calculated using Shannon entropy of MS/MS frequency distribution derived from the abundance of MS/MS precursors by the following equation as described by Martínez *et al.* (8).

$$H_j = -\sum_{i=1}^m P_{ij} \log_2(P_{ij})$$

where P_{ij} correspond to relative frequency of the i th MS/MS ($i = 1, 2, \dots, m$) in the j th sample ($j = 1, 2, \dots, t$).

The average frequency of the i th MS/MS among samples was calculated as

$$P_i = \frac{1}{t} \sum_{j=1}^t P_{ij}$$

Metabolic specificity, the S_i index, was defined as the expression identity of a given MS/MS regarding frequencies among considered samples. The MS/MS specificity was calculated as

$$S_i = \frac{1}{t} \left(\sum_{j=1}^t \frac{P_{ij}}{P_i} \log_2 \frac{P_{ij}}{P_i} \right)$$

The metabolome specialization δ_j index was measured for each j th sample, the average of the MS/MS specificities using the following formula

$$\delta_j = \sum_{i=1}^m P_{ij} S_i$$

MS/MS similarity scoring

MS/MS spectra were aligned in a pairwise manner and their similarity calculated according to two scores. First, a standard NDP, also referred to as cosine correlation method, was used to score fragment similarity among spectra using the following equation

$$\text{NDP} = \frac{\left(\sum_i^{S1 \& S2} W_{S1,i} W_{S2,i} \right)^2}{\sum_i W_{S1,i}^2 \sum_i W_{S2,i}^2}$$

where S1 and S2 correspond, respectively, to spectrum 1 and spectrum 2, and $W_{S1,i}$ and $W_{S2,i}$ indicate peak intensity-based weights given to i th common peaks differing by less than 0.01 Da between the two spectra. Weights were calculated as follows

$$W = [\text{Peak intensity}]^m [\text{Mass}]^n$$

with $m = 0.5$ and $n = 2$ as suggested by MassBank.

A second scoring method involving the analysis of shared NLS among individual MS/MS was implemented. For this, we used a list of 52 NLS commonly encountered during tandem MS fragmentation

and more specific ones that had been previously annotated for MS/MS spectra of *N. attenuata* secondary metabolite classes (data file S1) (9, 26). A binary vector of 1 and 0 was created for each MS/MS corresponding to present and absent of certain NL. NL similarity score was calculated for each pair of binary NL vectors based on Euclidean distance similarity.

MS/MS biclustering and molecular networking

To perform biclustering, we used the R package DiffCoEx, which is based on an extension of the weighted gene coexpression analysis (WGCNA). Using NDP and NL-scoring matrices for MS/MS spectra, we computed a comparative correlation matrix using DiffCoEx. The biclustering was performed with the parameters of “cutreeDynamic” set to method = “hybrid”, cutHeight = 0.9999, deepSplit = T, and minClusterSize = 10. The R source code of DiffCoEx is downloaded from additional file 1 by Tesson *et al.* (57); the required R WGCNA package can be found at <https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA>.

To perform MS/MS molecular networking analysis, we calculated pairwise spectral connectivity based on both NDP and NL similarity types and visualized the topology of the network using Cytoscape software with the organic layout in the yFiles layout algorithms extension app for Cytoscape.

Statistical analysis

Statistical analysis of the data was performed using R version 3.0.1. Statistical significance was evaluated using two-way analysis of variance (ANOVA), followed by Tukey’s honestly significant difference (HSD) post hoc tests. For analysis of differences between herbivory treatments and controls, Student’s *t* tests were used with the two-tailed distribution of two sets of samples with equal variance.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/24/eaaz0381/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- A. Kessler, A. Kalske, Plant secondary metabolite diversity and species interactions. *Annu. Rev. Ecol. Evol. Syst.* **49**, 115–138 (2018).
- L. A. Dyer, C. S. Philbin, K. M. Ochsenrider, L. A. Richards, T. J. Massad, A. M. Smilanich, M. L. Forister, T. L. Parchman, L. M. Galland, P. J. Hurtado, A. E. Espeset, A. E. Glassmire, J. G. Harrison, C. Mo, S. Yoon, N. A. Pardikes, N. D. Muchoney, J. P. Jahner, H. L. Slinn, O. Shelef, C. D. Dodson, M. J. Kato, L. F. Yamaguchi, C. S. Jeffrey, Modern approaches to study plant-insect interactions in chemical ecology. *Nat. Rev. Chem.* **2**, 50–64 (2018).
- P. W. Sherman, The levels of analysis. *Anim. Behav.* **36**, 616–619 (1988).
- M. C. Schuman, I. T. Baldwin, The layers of plant responses to insect herbivores. *Annu. Rev. Entomol.* **61**, 373–394 (2016).
- B. E. Sedio, Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification. *New Phytol.* **214**, 952–958 (2017).
- C. E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
- T. D. Schneider, D. N. Mastrorarde, Fast multiple alignment of ungapped DNA sequences using information theory and a relaxation method. *Discrete Appl. Math.* **71**, 259–268 (1996).
- O. Martínez, M. H. Reyes-Valdés, Defining diversity, specialization, and gene specificity in transcriptomes through information theory. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9709–9714 (2008).
- D. Li, S. Heiling, I. T. Baldwin, E. Gaquerel, Illuminating a plant’s tissue-specific metabolic diversity using computational metabolomics and information theory. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7610–E7618 (2016).
- D. F. Rhoades, in *Evolution of Plant Chemical Defense Against Herbivores* (Academic Press, 1979), pp. 1–55.
- F. R. Adler, R. Karban, Defended fortresses or moving targets? Another model of inducible defenses inspired by military metaphors. *Am. Nat.* **144**, 813–832 (1994).
- P. Feeny, in *Plant apparency and chemical defense. Recent Advances in Phytochemistry* (Biochemical Interaction Between Plants and Insects, Springer, 1976), vol. 10, pp. 1–40.
- J. P. Bryant, F. S. Chapin III, D. R. Klein, Carbon nutrient balance of boreal plants in relation to vertebrate herbivory. *Oikos* **40**, 357–368 (1983).
- P. D. Coley, J. P. Bryant, F. S. Chapin III, Resource availability and plant antiherbivore defense. *Science* **230**, 895–899 (1985).
- W. E. Loomis, Growth-differentiation balance vs carbohydrate-nitrogen ratio. *Proc. Am. Soc. Hortic. Sci.* **29**, 240–245 (1932).
- R. Karban, The ecology and evolution of induced responses to herbivory and how plants perceive risk. *Ecol. Entomol.* **45**, 1–9 (2020).
- A. C. McCall, J. A. Forgyce, Can optimal defence theory be used to predict the distribution of plant chemical defences? *J. Ecol.* **98**, 985–992 (2010).
- T. E. Ohmheiss, I. T. Baldwin, Optimal defense theory predicts the ontogeny of an induced nicotine defense. *Ecology* **81**, 1765–1783 (2000).
- I. T. Baldwin, Jasmonate-induced responses are costly but benefit plants under attack in native populations. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8113–8118 (1998).
- B. D. Moore, R. L. Andrew, C. Külheim, W. J. Foley, Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytol.* **201**, 733–750 (2014).
- L. A. Richards, L. A. Dyer, M. L. Forister, A. M. Smilanich, C. D. Dodson, M. D. Leonard, C. S. Jeffrey, Phytochemical diversity drives plant-insect community diversity. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10973–10978 (2015).
- G. S. Fraenkel, The raison d’être of secondary plant substances. *Science* **129**, 1466–1470 (1959).
- A. A. Agrawal, A. P. Hastings, M. T. J. Johnson, J. L. Maron, J.-P. Salminen, Insect herbivores drive real-time ecological and evolutionary change in plant populations. *Science* **338**, 113–116 (2012).
- A. Kempel, M. Schädler, T. Chrobok, M. Fischer, M. van Kleunen, Tradeoffs associated with constitutive and induced plant resistance against herbivory. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5685–5689 (2011).
- S. Y. Strauss, J. A. Rudgers, J. A. Lau, R. E. Irwin, Direct and ecological costs of resistance to herbivory. *Trends Ecol. Evol.* **17**, 278–285 (2002).
- D. Li, I. T. Baldwin, E. Gaquerel, Navigating natural variation in herbivory-induced secondary metabolism in coyote tobacco populations using MS/MS structural analysis. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4147–E4155 (2015).
- E. Gaquerel, M. Stitz, M. Kallenbach, I. T. Baldwin, Jasmonate signaling in the field, part II: Insect-guided characterization of genetic variations in jasmonate-dependent defenses of transgenic and natural *Nicotiana attenuata* populations. *Methods Mol. Biol.* **1011**, 97–109 (2013).
- P. B. Pelsler, H. de Vos, C. Theuring, T. Beuerle, K. Vrieling, T. Hartmann, Frequent gain and loss of pyrrolizidine alkaloids in the evolution of *Senecio* section *Jacobaea* (Asteraceae). *Phytochemistry* **66**, 1285–1295 (2005).
- T. Züst, C. Heichinger, U. Grossniklaus, R. Harrington, D. J. Kliebenstein, L. A. Turnbull, Natural enemies drive geographic variation in plant defenses. *Science* **338**, 116–119 (2012).
- F. Valladares, D. Sanchez-Gomez, M. A. Zavala, Quantitative estimation of phenotypic plasticity: Bridging the gap between the evolutionary concept and its ecological applications. *J. Ecol.* **94**, 1103–1116 (2006).
- C. Hettenhausen, I. T. Baldwin, J. Q. Wu, *Nicotiana attenuata* MPK4 suppresses a novel jasmonic acid (JA) signaling-independent defense pathway against the specialist insect *Manduca sexta*, but is not required for the resistance to the generalist *Spodoptera littoralis*. *New Phytol.* **199**, 787–799 (2013).
- H. Kaur, N. Heinzl, M. Schöttner, I. T. Baldwin, I. Gális, R2R3-NaMYB8 regulates the accumulation of phenylpropanoid-polyamine conjugates, which are essential for local and systemic defense against insect herbivores in *Nicotiana attenuata*. *Plant Physiol.* **152**, 1731–1747 (2010).
- S. Heiling, M. C. Schuman, M. Schoettner, P. Mukerjee, B. Berger, B. Schneider, A. R. Jassbi, I. T. Baldwin, Jasmonate and ppHsystemin regulate key malonylation steps in the biosynthesis of 17-hydroxygeranylinalool diterpene glycosides, an abundant and effective direct defense against herbivores in *Nicotiana attenuata*. *Plant Cell* **22**, 273–292 (2010).
- E. S. McCloud, I. T. Baldwin, Herbivory and caterpillar regurgitants amplify the wound-induced increases in jasmonic acid but not nicotine in *Nicotiana sylvestris*. *Planta* **203**, 430–435 (1997).
- S.-G. Kim, F. Yon, E. Gaquerel, J. Gulati, I. T. Baldwin, Tissue specific diurnal rhythms of metabolites and their regulation during herbivore attack in a native tobacco, *Nicotiana attenuata*. *PLOS ONE* **6**, e26214 (2011).
- M. Kallenbach, G. Bonaventure, P. A. Gilardoni, A. Wissgott, I. T. Baldwin, *Empoasca* leafhoppers attack wild tobacco plants in a jasmonate-dependent manner and identify jasmonate mutants in natural populations. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1548–E1557 (2012).
- T. H. Goodspeed, *The genus Nicotiana* (Chronica Botanica Company, Waltham, Mass., 1954), vol. 16, pp. 102–35.
- J. J. Clarkson, L. J. Kelly, A. R. Leitch, S. Knapp, M. W. Chase, Nuclear glutamine synthetase evolution in *Nicotiana*: Phylogenetics and the origins of allotetraploid and homoploid (diploid) hybrids. *Mol. Phylogenet. Evol.* **55**, 99–112 (2010).

39. W. Zhou, T. Brockmüller, Z. Ling, A. Omdahl, I. T. Baldwin, S. Xu, Evolution of herbivore-induced early defense signaling was shaped by genome wide duplications in *Nicotiana glauca*. *eLife* **5**, e19531 (2016).
40. S. Xu, W. Zhou, S. Pottinger, I. T. Baldwin, Herbivore associated elicitor-induced defences are highly specific among closely related *Nicotiana* species. *BMC Plant Biol.* **15**, 2 (2015).
41. J. R. Platt, Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science* **146**, 347–353 (1964).
42. N. Stamp, Out of the quagmire of plant defense hypotheses. *Q. Rev. Biol.* **78**, 23–55 (2003).
43. R. E. Ulanowicz, Information theory in ecology. *Comput. Chem.* **25**, 393–399 (2001).
44. U. Schittko, C. A. Preston, I. T. Baldwin, Eating the evidence? *Manduca sexta* larvae can disrupt specific jasmonate induction in *Nicotiana attenuata* by rapid consumption. *Planta* **210**, 343–346 (2000).
45. H. Guo, N. Wielsch, J. B. Hafke, A. Svatoš, A. Mithöfer, W. Boland, A porin-like protein from oral secretions of *Spodoptera littoralis* larvae induces defense-related early events in plant leaves. *Insect Biochem Mol. Biol.* **43**, 849–858 (2013).
46. L. Mack, P. Gros, J. Burkhardt, K. Seifert, Elicitors of tansy volatiles from cotton leafworm larval oral secretion. *Phytochemistry* **96**, 158–169 (2013).
47. R. Schweiger, A.-M. Heise, M. Persicke, C. Müller, Interactions between the jasmonic and salicylic acid pathway modulate the plant metabolome and affect herbivores of different feeding types. *Plant Cell Environ.* **37**, 1574–1585 (2014).
48. C. C. von Dahl, R. A. Winz, R. Halitschke, F. Kühnemann, K. Gase, I. T. Baldwin, Tuning the herbivore-induced ethylene burst: The role of transcript accumulation and ethylene perception in *Nicotiana attenuata*. *Plant J.* **51**, 293–307 (2007).
49. R. A. Winz, I. T. Baldwin, Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. IV. Insect-induced ethylene reduces jasmonate-induced nicotine accumulation by regulating putrescine *N*-methyltransferase transcripts. *Plant Physiol.* **125**, 2189–2202 (2001).
50. R. H. Whittaker, Evolution and measurement of species diversity. *Taxon* **21**, 213–251 (1972).
51. P. R. Ehrlich, P. H. Raven, Butterflies and plants: A study in coevolution. *Evolution* **18**, 586–608 (1964).
52. A. A. Agrawal, M. Fishbein, R. Halitschke, A. P. Hastings, D. L. Rabosky, S. Rasmann, Evidence for adaptive radiation from a phylogenetic study of plant defenses. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18067–18072 (2009).
53. R. D. Finn, C. G. Jones, Natural products—A simple model to explain chemical diversity. *Nat. Prod. Rep.* **20**, 382–391 (2003).
54. M. Schäfer, C. Brütting, I. T. Baldwin, M. Kallenbach, High-throughput quantification of more than 100 primary- and secondary-metabolites, and phytohormones by a single solid-phase extraction based sample preparation with analysis by UHPLC-HESI-MS/MS. *Plant Methods* **12**, (2016).
55. A. Paschold, R. Halitschke, I. T. Baldwin, Co(i)-ordinating defenses: NaCO11 mediates herbivore-induced resistance in *Nicotiana attenuata* and reveals the role of herbivore movement in avoiding defenses. *Plant J.* **51**, 79–91 (2007).
56. M. Perlea, D. Kim, G. M. Perlea, J. T. Leek, S. L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
57. B. M. Tesson, R. Breitling, R. C. Jansen, DiffCoEx: A simple and sensitive method to find differentially coexpressed gene modules. *BMC Bioinf.* **11**, 497 (2010).

Acknowledgments: We thank W. Zhou for helping with the samples, R. Ray for the support in assembling the RNA-seq dataset, and H. Guo and Z. Jaramillo for the help in establishing ET measurement. **Funding:** D.L., R.H., and I.T.B. were funded by the Max-Planck-Society, an advanced grant no. 293926 of the European Research Council to I.T.B. and the Collaborative Research Centre “Chemical Mediators in Complex Biosystems-ChemBioSys” (SFB 1127). E.G.’s research was supported within the framework of the Deutsche Forschungsgemeinschaft Excellence Initiative to the University of Heidelberg and by the CNRS in Strasbourg. **Author contributions:** D.L. conceived the study, performed the experiments, analyzed the data, and wrote the paper. R.H. helped conceive the study. E.G. and I.T.B. conceived the study, analyzed the data, and wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** MS/MS data are deposited in the European Molecular Biology Laboratory European Bioinformatics Institute open metabolomics database MetaboLights (www.ebi.ac.uk; under the accession no. MTBLS1471). Microarray data are deposited in NCBI Gene Expression Omnibus database under the accession no. GSE30287. All raw RNA-seq data of *Nicotiana* species are accessible under NCBI SRA under the project number PRJNA301787. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 7 August 2019

Accepted 14 April 2020

Published 10 June 2020

10.1126/sciadv.aaz0381

Citation: D. Li, R. Halitschke, I. T. Baldwin, E. Gaquerel, Information theory tests critical predictions of plant defense theory for specialized metabolism. *Sci. Adv.* **6**, eaaz0381 (2020).