Tansley insight

Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification

Author for correspondence:
Brian E. Sedio
Tel: +507 212 8763
Email: SedioB@si.edu

Received: *7 October 2016*
Accepted: *2 December 2016*

Brian E. Sedio[1,2]

[1]Smithsonian Tropical Research Institute, Apartado 0843–03092, Balboa, Ancón, Republic of Panama; [2]Center for Biodiversity and Drug Discovery, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología, Apartado 0843-01103, Ciudad del Saber, Ancón, Republic of Panama

Contents

New Phytologist (2017) **214:** 952–958
doi: 10.1111/nph.14438

Key words: chemical evolution, community ecology, macroevolution, mass spectrometry, nuclear magnetic resonance spectroscopy, species coexistence.

Summary

Much of our understanding of the mechanisms by which biotic interactions shape plant communities has been constrained by the methods available to study the diverse secondary chemistry that defines plant relationships with other organisms. Recent innovations in analytical chemistry and bioinformatics promise to reveal the cryptic chemical traits that mediate plant ecology and evolution by facilitating simultaneous structural comparisons of hundreds of unknown molecules to each other and to libraries of known compounds. Here, I explore the potential for mass spectrometry and nuclear magnetic resonance metabolomics to enable unprecedented tests of seminal, but largely untested hypotheses that propose a fundamental role for plant chemical defenses against herbivores and pathogens in the evolutionary origins and ecological coexistence of plant species diversity.

I. Introduction

Seminal hypotheses in community ecology and evolutionary biology propose a fundamental role for plant chemical defenses against herbivores and pathogens in the evolutionary origins and ecological coexistence of plant species. Whereas nearly all plants require a small number of shared resources (water, light, CO_2 and nutrients), plant interactions with natural enemies provide a highly multidimensional space within which species can carve out a distinct niche defined by the enemies they support, and those they avoid. Gillett (1962) first proposed that plants build up host-specific natural enemies where they are abundant, impeding their fitness relative to competitors that manage to avoid sharing the enemy. Janzen (1970) and Connell (1971) proposed that the resulting negative density-dependent recruitment is a primary driver of plant species coexistence in tropical forests. Ehrlich & Raven (1964) extended the concept to macroevolution, proposing that the ecological success conferred by novel defenses against natural enemies facilitates speciation, and hence lineage diversification, in plants (and in their enemies; see Box 1).

Box 1 Glossary of key terms

Coexistence: the stable co-occurrence of species resulting from niche differentiation.

Competitive exclusion: the elimination from a community of one of two species with unequal fitness and overlapping resource requirements or natural enemies. Shared natural enemies give rise to 'apparent competitive exclusion' (Holt, 1977).

Key innovation: a novel character state that is associated with an increase in the rate of phylogenetic lineage diversification.

Mass spectrum: a plot of ion intensity vs mass-to-charge ratio (m/z) of molecules or molecular fragments.

Moiety: part of a molecule (e.g. a ketone moiety consists of a carbon atom with a double-bond to an oxygen atom and two single bonds to other carbon atoms).

Niche differentiation: species differences in resource requirements or defenses such that competition for resources or the likelihood of attack by enemies is greater when neighbors are conspecifics.

Nuclear magnetic resonance (NMR): the absorption and re-emission of electromagnetic radiation by atomic nuclei, dependent upon the magnetic fields of nearby atoms and therefore indicative of molecular structure.

Rate of character evolution: the rate over time or phylogenetic branch length at which a trait changes over a phylogenetic tree; can be defined in terms of trait contrasts between sister lineages in a phylogeny.

Rate of diversification: the rate of accumulation of species or evolutionary lineages over time; speciation minus extinction.

Scaffold: molecular backbone used to classify compounds into broad classes (e.g. a benzene ring).

Many thousands of secondary metabolites influence interactions between plants and herbivores and pathogens. For example, the bean family (Fabaceae) alone synthesizes thousands of compounds from nearly 20 major chemical classes (Wink & Mohamed, 2003). The sheer number of secondary metabolites of unknown structure has long precluded comparative metabolomics, the comparison of small-molecule metabolite profiles, at the large taxonomic scales required for the study of macroevolution and community ecology (Hilker, 2014). However, recent advances in tandem mass spectrometry (MS/MS) and nuclear magnetic resonance (NMR) spectroscopy bioinformatics make chemical analysis at community and macroevolutionary scales possible by enabling comparison of the structure of unknown compounds from crude extracts of chemically complex biological samples.

Here, I illustrate the potential for MS/MS and NMR structural metabolomics to permit unprecedented tests of the hypotheses that plant secondary chemistry (1) determines host ranges of pathogens and herbivores and thereby (2) facilitates species coexistence in hyperdiverse plant communities such as tropical forest tree communities and (3) promotes the evolutionary diversification of plants due to the fitness advantage conferred by the evolution of novel chemical defenses.

II. Recent innovations in structural metabolomics

Until recently, untargeted metabolomic profiling of chemically complex samples has required the isolation of individual compounds for manual structural determination. By contrast, innovations in bioinformatics use NMR and MS/MS to compare the structures of unknown compounds (Table 1). Coupled, automated separation and detection methods, such as liquid chromatography (LC)-solid-phase extraction (SPE)-MS/NMR, have enhanced the efficiency of the collection of MS and NMR spectra from chemically complex samples (Moco & Vervoort, 2012). The interpretation of NMR spectra allows unequivocal determination of molecular structure, which allows compounds to be classified based on their structural scaffold or by metabolic pathway in well-studied organisms (Wetzel *et al.*, 2007). However, the isolation of individual compounds of interest, such as with SPE, is a rate- and cost-limiting step even when automated at micro-volumes. For these reasons, applications of metabolomics to understanding the role of plant secondary chemistry in community ecology and macroevolution have been limited. Recent advances in both MS/MS and NMR bioinformatics, however, make high-throughput chemical analysis at the scale of a species-rich ecological community or phylogenetic clade possible by quantifying the structural similarity of samples that are complex mixtures of compounds of unknown structure.

Recent innovations in NMR bioinformatics enable the quantification of molecular structural diversity in complex samples of unknown compounds (Richards *et al.*, 2015). The NMR method begins with ^{1}H-NMR spectra collected from crude tissue extracts. The diversity of ^{1}H-NMR resonances in a spectrum is indicative of molecular structural (scaffold and moiety) diversity. In pairwise comparisons of NMR spectra, an unique ^{1}H-NMR resonance indicates the presence of a unique moiety in one of the samples, and the similarity of crude-extract spectra for two plant species can be interpreted as the chemical structural similarity of their metabolomes (Table 1).

MS/MS enables comparisons of the chemical structure of compounds because molecules with similar structures fragment into many of the same sub-structures (Fig. 1a,b). Thus, structural similarity can be quantified for thousands of pairs of unknown compounds by measuring the cosine of the angle between vectors that represent the mass-to-charge ratio (m/z) of the constituent

Table 1 Comparison of three alternative methods for untargeted metabolomics

Method	Unequivocal identification of unknown compounds	Method of identification of unknown compounds	Metric of molecular structural similarity	Metric of sample structural similarity/diversity	Relative throughput	References
LC-SPE-MS/NMR target identification	Yes	Manual analysis of NMR spectra	Scaffold-based classification, metabolic classification	Chemical structural-compositional similarity (CSCS)	Low	Wetzel *et al.* (2007); Moco & Vervoort (2012)
Crude-extract NMR	No	^{1}H-NMR match to annotated spectra of compound in isolation	Scaffold-based classification	Chemical diversity index (CDI)	Medium	Richards *et al.* (2015)
LC-MS/MS networking	No	MS/MS match to annotated spectra of known compound	Cosine of MS/MS spectra	Chemical structural-compositional similarity (CSCS)	High	Watrous *et al.* (2012); Wang *et al.* (2016); Sedio *et al.* (in press)

Unequivocal identification of unknown compounds refers to the capacity to determine the structure of a novel metabolite using MS and NMR spectra without the aid of reference, annotated spectra. Method of identification of unknown compounds refers to 'dereplication,' the confirmation of the structure of a metabolite from a complex mixture by comparison with an annotated MS or NMR spectrum of the known compound. The structural similarity of known compounds can be quantified by classification with respect to molecular scaffold or the metabolic pathway from which they are derived, if it is known. The structural similarity of unknown compounds can be quantified by calculating the cosine of pairs of MS/MS spectra. Richards *et al.* (2015) and Sedio *et al.* (in press) describe CDI and CSCS metrics, respectively, for quantifying the structural diversity or similarity of complex mixtures. The capacity to derive metabolomics data from complex mixtures without isolating or identifying compounds makes crude-extract NMR and especially MS/MS molecular networking massively scalable. LC, liquid chromatography; SPE, solid-phase extraction; MS, mass spectrometry; MS/MS, tandem MS; NMR, nuclear magnetic resonance.

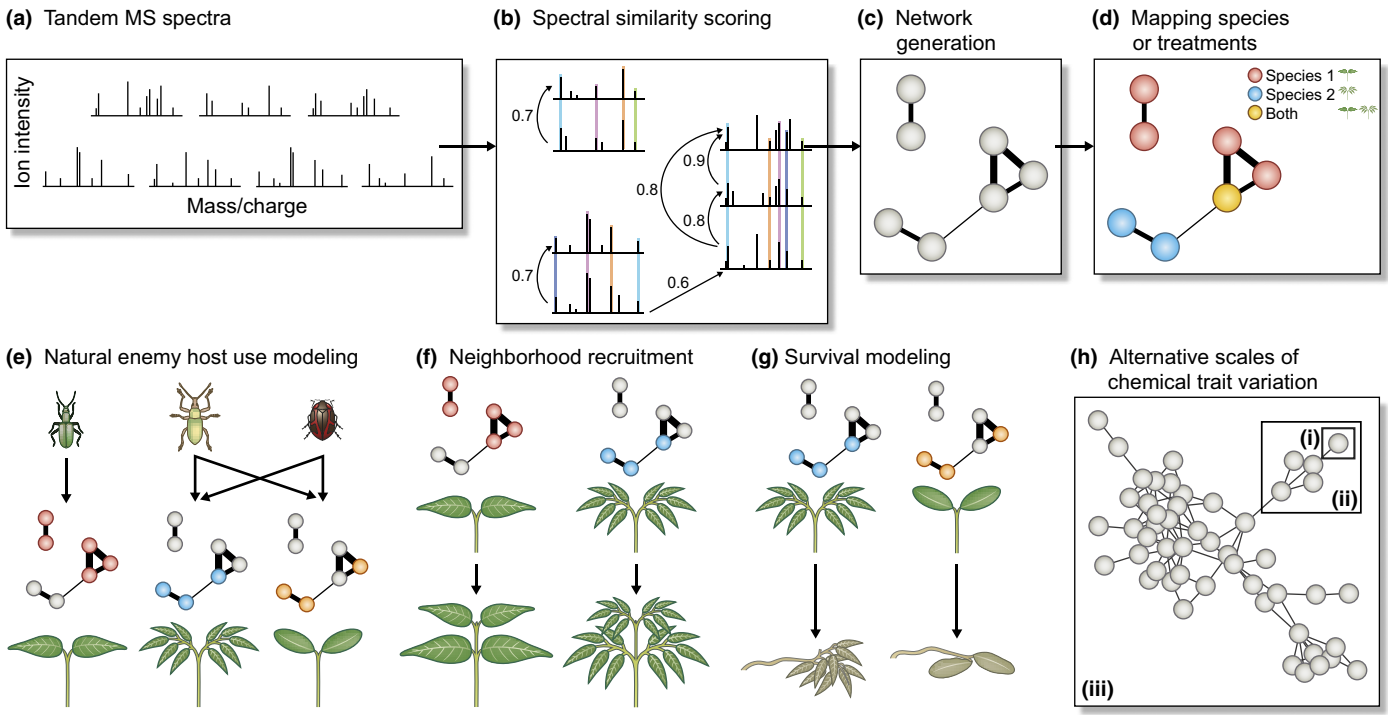

Fig. 1 The application of tandem mass spectrometry (MS/MS) to plant-enemy ecology. MS/MS of crude plant extracts provides molecular spectra representing seven compounds, with each peak representing the mass-to-charge ratio (*m/z*, horizontal axis) and ion intensity (vertical axis) of a constituent molecular fragment (a). Spectra are aligned (colored vertical lines identify shared fragments) and pairwise similarity scores (arrows and numbers) are calculated (b). The similarity scores define molecular networks in which nodes represent compounds and the width of the links represent pairwise structural similarity (c). Compounds are mapped onto two plant species (d). Herbivore specificity and host chemical similarity are related in a three plant species example; colors indicate compound incidence in species; compounds not found in each species are gray (e). Neighboring plants prosper if they are chemically dissimilar (f) but suffer attack by shared herbivores and suffer local mortality if they are chemically similar (g). In a molecular network of compounds linked by structural similarity, boxes illustrate three, alternative scales of molecular structural variation at which host use (e) and recruitment (f, g) models might consider chemical trait variation among plant species: (i) single compounds, (ii) small clusters of highly structurally similar compounds that may be derived from a common metabolic pathway, and (iii) large subnetworks of compounds with common structural features that may represent chemical classes (h).

fragments (Table 1; Watrous *et al.*, 2012). In addition, comparison of MS/MS spectra with publically available spectral libraries can identify unknown molecules ('dereplication'; Allard *et al.*, 2016; Wang *et al.*, 2016). More importantly, this method enables the quantification of molecular similarity even for samples in which few compounds are unambiguously identified, permitting chemical ecology in understudied and species-rich plant communities such as tropical forests.

The structural comparison of unknown molecules using MS/MS is scalable to datasets containing hundreds of samples and tens of thousands of unique molecules (Watrous *et al.*, 2012; Wang *et al.*, 2016). Visualization and analysis is aided by the organization of pairwise MS/MS similarities into molecular networks, in which nodes represent compounds and links between nodes represent similarity scores (Fig. 1c). Furthermore, molecular networks aid in the classification of unknown molecules that are linked to compounds that match annotated spectra in public libraries (Wang *et al.*, 2016).

The vast diversity of plant secondary metabolites and the rarity of any particular compound among species have long stymied ecologists and evolutionary biologists. Structural metabolomic methods address these problems by accounting for the structural similarity of unknown compounds that are not shared between species when quantifying chemical similarity (Table 1). Richards *et al.* (2015) developed a chemical diversity index (CDI) based on crude ^{1}H-NMR spectra that reflects both inter- and intramolecular moiety diversity. Sedio *et al.* (in press) developed a chemical structural-compositional similarity (CSCS) metric that weights the structural similarity of every pair of compounds in a network by their relative ion intensity in two plant species. Conventional methods of calculating similarity in ecology, such as Bray–Curtis similarity, consider the relative abundance of shared compounds, but ignore structural relationships between molecules. By contrast, both CDI and CSCS account for the presence of structurally similar compounds that are not shared between species or samples. A simple example illustrates the implications. Compounds *x* and *y* are structurally similar, species *A* contains compound *x* but not *y*, and species *B* contains *y* but not *x*. In this example, compounds *x* and *y* make no contribution to Bray–Curtis similarity, but make a positive contribution to CDI and CSCS.

Furthermore, the proximity of compounds in an MS/MS molecular network can be used to quantify structural scale, from pairs of highly structurally similar compounds that share a direct link, to subnetworks of compounds with shared structural scaffolds, to large clusters that may correspond to chemical classes with shared structural elements (Fig. 1h). By enabling the high-throughput structural comparison of unknown molecules, structural metabolomics promises unprecedented insight into the secondary-chemistry niches hypothesized to generate and sustain plant diversity.

III. Species coexistence

Can the number of niches defined by secondary metabolites and their impact on plant enemies approach the number of coexisting tree species in a tropical forest? Or, is much of the variation in secondary chemistry and other defenses redundant in the eyes of plant enemies? Biologists have accumulated 40 yr of evidence in support of the predictions of Gillett (1962), Janzen (1970) and Connell (1971) that seeds and seedlings experience reduced survival and recruitment in the vicinity of conspecific adults (reviewed in Comita *et al.*, 2014). Density-dependent suppression of conspecific individuals suggests that natural enemy host ranges are sufficiently narrow to ensure that enemy-mediated competition (Holt, 1977) is greater among conspecific than between heterospecific individuals, the definition of niche differentiation and a prerequisite for coexistence (Chesson & Kuang, 2008).

However, many herbivores (Novotny *et al.*, 2002) and pathogens (Gilbert & Webb, 2007) are not strict specialists. Generalist plant enemies, even those with narrow host ranges, are expected to mediate competitive exclusion of shared hosts, effectively limiting the co-occurrence of species that are chemically similar and promoting chemical diversity in the plant community (Fig. 1f,g; Sedio & Ostling, 2013).

Secondary chemistry exhibits phylogenetic signal at broad phylogenetic scales (Wink, 2003). However, secondary chemistry is often not conserved within species-rich genera, and closely related species can differ dramatically chemically in *Bursera* (Becerra, 1997), *Inga* (Kursar *et al.*, 2009), *Protium* (Fine *et al.*, 2013), *Solanum* (Haak *et al.*, 2014), *Piper* (Richards *et al.*, 2015; Salazar *et al.*, 2016), *Eugenia*, *Ocotea* and *Psychotria* (Sedio *et al.*, in press). Furthermore, co-occurring species of *Bursera* (Becerra, 2007), *Inga* (Kursar *et al.*, 2009) and *Piper* (Salazar *et al.*, 2016) are less similar chemically than by chance, suggesting that niche partitioning based on defense compounds stabilizes coexistence among species in these genera.

Structural metabolomic methods make it possible to move beyond individual genera to study chemical ecology at the scale of whole communities. In addition, recent developments in DNA barcoding (Garcia-Robledo *et al.*, 2013) and microbial metagenomics (Barberán *et al.*, 2015) enable determination of plant–insect and plant–microbe associations at large community scales. The standardized application of metabolomics and metagenomics across multiple sites can facilitate climatic, latitudinal and biogeographical comparisons. To implicate herbivores and/or pathogens in plant species coexistence, one might statistically infer the relative explanatory power of chemical traits in determining host use patterns of natural enemies (Fig. 1e), and similarly, infer the power of local neighborhood densities of those traits (or mutual natural enemies themselves) in determining plant performance and recruitment (Fig. 1f,g).

Structural metabolomic data can be used to infer the chemical traits that determine plant–enemy associations by modeling natural enemy host use as a multinomial distribution over potential host plants, with a probability vector that is a function of plant chemistry (Fig. 1e). Alternatively, models used in genome-wide association studies to identify loci associated with discrete phenotypes out of thousands of candidate loci could be modified to instead identify compounds (analogous to loci) associated with the presence or absence of particular natural enemies (analogous to phenotypes) out of thousands of candidate compounds. Similarly, the prediction that certain chemical traits are associated with

density-dependent neighborhood effects on recruitment or mortality (Fig. 1f,g) can be tested by drawing on a deep literature relating traits to density-dependent dynamics and community structure (e.g. Kraft *et al.*, 2008; Comita *et al.*, 2010; Pollock *et al.*, 2012; Sedio & Ostling, 2013; Lebrija-Trejos *et al.*, 2016; Salazar *et al.*, 2016).

Alternative models of natural enemy host associations or of density-dependent performance might consider chemical 'traits' that represent MS/MS subnetworks or NMR functional groups at various degrees of inclusiveness, from scales at which groups represent broad compound classes (e.g. xanthine-derivative alkaloids) to much narrower groups of structurally similar compounds (e.g. caffeine and theobromine), as illustrated in Fig. 1(h). Such model comparisons could reveal the scale at which chemical structural variation influences herbivore and microbe host associations and at what scale defensive compounds are functionally redundant.

IV. Character evolution and lineage diversification

Ehrlich & Raven (1964) first proposed that diversification in plants is often the result of innovation in defenses against natural enemies. Their hypothesis is referred to as the 'Escape and Radiate' Hypothesis because it envisions the evolution of a novel defense (e.g. the two-compound cluster in Fig. 2a), subsequent ecological success as the plant population grows unchecked by natural enemies, followed by speciation, and ultimately diversification of many species descended from the original plant species (Fig. 2b; Schluter, 2000). Two mutually compatible predictions follow. Species richness or the rate of diversification (speciation minus extinction) in phylogenetic clades should be associated with the evolution of key innovations in defense. And, variation in rates of lineage diversification over a plant phylogeny should be associated with variation in rates of defense evolution (Fig. 2b).

There have been two tests of the key innovation hypothesis. Farrell *et al.* (1991) demonstrated that plant lineages that exude latex from damaged tissue exhibit greater species richness than sister lineages that lack latex, suggesting that latex is a key innovation associated with adaptive radiations in plants. Similarly, Weber & Agrawal (2014) found that plant lineages in which some species use extra-floral nectaries to recruit ants to their defense show greater rates of diversification than sister lineages that lack nectaries. Latex and extra-floral nectaries are recognizable characters in distantly related lineages of plants. Secondary chemistry has proved a more challenging subject for macroevolutionary analyses. The phylogenetic rarity of any particular compound makes it difficult to identify potential key innovations. Likewise, the astonishing diversity of compounds that plants deploy as defenses makes it difficult to identify comparable characters to compare rates of character evolution across plant lineages. For these reasons, the predictions of Ehrlich & Raven (1964) have remained largely untested with respect to secondary chemistry at taxonomic scales beyond closely related congeners (e.g. Agrawal *et al.*, 2009).

Structural metabolomics promises to open a new frontier in the study of chemical macroevolution by providing a common currency by which to measure character evolution of compounds of unknown structure in distantly related lineages. The hyperdiverse tree genera *Inga, Piper* and *Psychotria* deploy distinct chemical classes in defense (Kursar *et al.*, 2009; Richards *et al.*, 2015; Sedio *et al.*, in press). In an MS/MS molecular network or an NMR metabolic profile, interspecific variation in unknown phenolic compounds of *Inga*, sesqui- and tri-terpenes of *Piper*, and indole

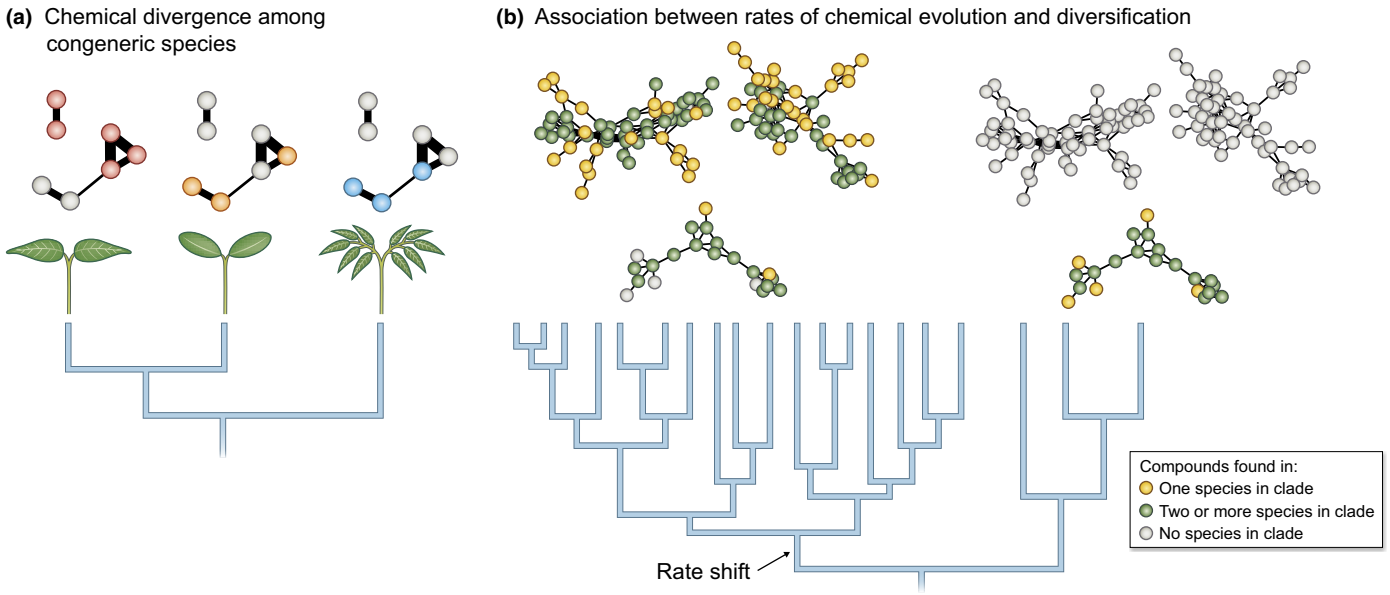

Fig. 2 Structural metabolomics provides a common currency to measure variation in rates of chemical evolution in distantly related phylogenetic lineages. Congeneric species in the seven species-rich genera that have been studied exhibit a conspicuous absence of phylogenetic signal (a, where the two closely related species on the left are more chemically distinct than are the two distantly related species on the right). Mass spectrometry molecular networks provide comparable chemical trait data for distantly related plant lineages in which distinct chemical classes predominate, making possible tests of Ehrlich & Raven's (1964) seminal hypothesis that defense evolution drives lineage diversification (b).

and pyridone alkaloids of *Psychotria* are quantified on a comparable scale (e.g. using CDI or CSCS; Richards *et al.*, 2015; Sedio *et al.*, in press).

Rates of chemical evolution can easily be quantified by measuring phylogenetic independent contrasts (PICs; Felsenstein, 1985) in terms of CSCS (Sedio *et al.*, in press) for every branch in a phylogeny. Bayesian phylogenetic comparative methods, such as the Bayesian Analysis of Macroevolutionary Mixtures (BAMM; Rabosky, 2014), can identify the location and number of key shifts in the rate of diversification or character evolution without *a priori* predictions. Because MS/MS data are increasingly shared through repositories such as the Global Natural Products Social (GNPS) Molecular Networking database (Wang *et al.*, 2016), the availability of metabolomic data and appropriate phylogenetic comparative methods will allow unprecedented tests of the role of secondary chemistry in the evolutionary origins of plant diversity at global scales.

V. Conclusions

Seminal hypotheses that interspecific variation in secondary chemistry enables species coexistence and drives evolutionary diversification in plants have remained largely untested due to an inability to measure chemical traits comprehensively at scales appropriate for studies of community ecology and macroevolution. Advances in MS/MS (Watrous *et al.*, 2012; Wang *et al.*, 2016) and NMR (Richards *et al.*, 2015) bioinformatics enable comparative structural metabolomics, that is, the ability to quantify the structural similarity of thousands of unknown secondary compounds in hundreds of species at a time. The tools now exist to reveal the cryptic chemical traits that were hypothesized to drive global patterns of diversity among communities and among evolutionary lineages of plants more than half a century ago (Gillett, 1962; Ehrlich & Raven, 1964; Janzen, 1970; Connell, 1971).

Acknowledgements

I thank Pieter C. Dorrestein, E. Allen Herre and S. Joseph Wright for helpful comments and discussions. This work was supported by the Smithsonian Tropical Research Institute Earl S. Tupper Fellowship and the Smithsonian Institution Scholarly Studies Awards Program and Grand Challenges Consortium.

References

Agrawal AA, Fishbein M, Halitschke R, Hastings AP, Rabosky DL, Rasmann S. 2009. Evidence for adaptive radiation from a phylogenetic study of plant defenses. *Proceedings of the National Academy of Sciences, USA* **106**: 18067–18072.

Allard PM, Péresse T, Bisson J, Gindro K, Marcourt L, Pham VC, Roussi F, Litaudon M, Wolfender JL. 2016. Integration of molecular networking and *in-silico* MS/MS fragmentation for natural products dereplication. *Analytical Chemistry*, **88**: 3317–3323.

Barberán A, McGuire KL, Wolf JA, Jones FA, Wright SJ, Turner BL, Essene A, Hubbell SP, Faircloth BC, Fierer N. 2015. Relating belowground microbial composition to the taxonomic, phylogenetic, and functional trait distributions of trees in a tropical forest. *Ecology Letters* **18**: 1397–1405.

Becerra JX. 1997. Insects on plants: macroevolutionary chemical trends in host use. *Science* **276**: 253–256.

Becerra JX. 2007. The impact of herbivore-plant coevolution on plant community structure. *Proceedings of the National Academy of Sciences, USA* **104**: 7483–7488.

Chesson P, Kuang JJ. 2008. The interaction between predation and competition. *Nature* **456**: 235–238.

Comita LS, Muller-Landau HC, Aguilar S, Hubbell SP. 2010. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**: 330–332.

Comita LS, Queenborough SA, Murphy SJ, Eck JL, Xu K, Krishnadas M, Beckman N, Zhu Y. 2014. Testing predictions of the Janzen-Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *Journal of Ecology* **102**: 845–856.

Connell JH. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and rain forest trees. In: Boer PJD, Gradwell GR, eds. *Dynamics of populations*. Wageningen, the Netherlands: Center for Agricultural Publication and Documentation, 298–312.

Ehrlich PR, Raven PH. 1964. Butterflies and plants – a study in coevolution. *Evolution* **18**: 586–608.

Farrell BD, Dussourd DE, Mitter C. 1991. Escalation of plant defense – do latex and resin canals spur plant diversification. *American Naturalist* **138**: 881–900.

Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* **125**: 1–15.

Fine PVA, Metz MR, Lokvam J, Mesones I, Zuniga JMA, Lamarre GPA, Pilco MV, Baraloto C. 2013. Insect herbivores, chemical innovation, and the evolution of habitat specialization in Amazonian trees. *Ecology* **94**: 1764–1775.

Garcia-Robledo C, Erickson DL, Staines CL, Erwin TL, Kress WJ. 2013. Tropical plant-herbivore networks: reconstructing species interactions using DNA barcodes. *PLoS ONE* **8**: e52967.

Gilbert GS, Webb CO. 2007. Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Sciences, USA* **104**: 4979–4983.

Gillett JD. 1962. Pest pressure, an underestimated factor in evolution. *Systematics Association Publication* **4**: 37–46.

Haak DC, Ballenger BA, Moyle LC. 2014. No evidence for phylogenetic constraint on natural defense evolution among wild tomatoes. *Ecology* **95**: 1633–1641.

Hilker M. 2014. New synthesis: parallels between biodiversity and chemodiversity. *Journal of Chemical Ecology* **40**: 225–226.

Holt RD. 1977. Predation, apparent competition, and structure of prey communities. *Theoretical Population Biology* **12**: 197–229.

Janzen DH. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* **104**: 501–528.

Kraft NJB, Valencia R, Ackerly DD. 2008. Functional traits and niche-based tree community assembly in an amazonian forest. *Science* **322**: 580–582.

Kursar TA, Dexter KG, Lokvam J, Pennington RT, Richardson JE, Weber MG, Murakami ET, Drake C, McGregor R, Coley PD. 2009. The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proceedings of the National Academy of Sciences, USA* **106**: 18073–18078.

Lebrija-Trejos E, Reich PB, Hernandez A, Wright SJ. 2016. Species with greater seed mass are more tolerant of conspecific neighbours: a key driver of early survival and future abundances in a tropical forest. *Ecology Letters* **19**: 1071–1080.

Moco S, Vervoort J. 2012. Chemical identification strategies using liquid chromatography-photodiode array-solid-phase extraction-nuclear magnetic resonance/mass spectrometry. In: Hardy NW, Hall RD, eds. *Plant metabolomics: methods and protocols*. New York, NY, USA: Springer, 287–316.

Novotny V, Basset Y, Miller SE, Weiblen GD, Bremer B, Cizek L, Drozd P. 2002. Low host specificity of herbivorous insects in a tropical forest. *Nature* **416**: 841–844.

Pollock LJ, Morris WK, Vesk PA. 2012. The role of functional traits in species distributions revealed through a hierarchical model. *Ecography* **35**: 716–725.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**: e89543.

Richards LA, Dyer LA, Forister ML, Smilanich AM, Dodson CD, Leonard MD, Jeffrey CS. 2015. Phytochemical diversity drives plant–insect community diversity. *Proceedings of the National Academy of Sciences, USA* **112**: 10973–10978.

Salazar D, Jaramillo A, Marquis RJ. 2016. The impact of plant chemical diversity on plant–herbivore interactions at the community level. *Oecologia* **181**: 1199–1208.

Schluter D. 2000. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.

Sedio BE, Ostling AM. 2013. How specialised must natural enemies be to facilitate coexistence among plants? *Ecology Letters* **16**: 995–1003.

Sedio BE, Rojas Echeverri JC, Boya CA, Wright SJ. In press. Sources of variation in foliar secondary chemistry in a tropical forest tree community. *Ecology*. doi: 10.1002/ecy.1689.

Wang MX, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T *et al.* 2016. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology* **34**: 828–837.

Watrous J, Roach P, Alexandrov T, Heath BS, Yang JY, Kersten RD, van der Voort M, Pogliano K, Gross H, Raaijmakers JM *et al.* 2012. Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences, USA* **109**: E1743–E1752.

Weber MG, Agrawal AA. 2014. Defense mutualisms enhance plant diversification. *Proceedings of the National Academy of Sciences, USA* **111**: 16442–16 447.

Wetzel S, Schuffenhauer A, Roggo S, Ertl P, Waldmann H. 2007. Chemoinformatic analysis of natural products and their chemical space. *CHIMIA International Journal for Chemistry* **61**: 355–360.

Wink M. 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**: 3–19.

Wink M, Mohamed GIA. 2003. Evolution of chemical defense traits in the Leguminosae: mapping of distribution patterns of secondary metabolites on a molecular phylogeny inferred from nucleotide sequences of the rbcL gene. *Biochemical Systematics and Ecology* **31**: 897–917.