University of Nevada, Reno

Butterflies, inchworms, and plants: revisiting long-standing hypotheses of codiversification

A dissertation submitted in partial fulfillment of the
Requirements for the degree of Doctor of Philosophy in
Ecology, Evolution, and Conservation Biology

by

Joshua P. Jahner

Dr. Matthew L. Forister/Dissertation Advisor

August, 2015

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

JOSHUA P. JAHNER

Entitled

Butterflies, inchworms, and plants: revisiting long-standing hypotheses of codiversification.

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Matthew L. Forister, Advisor

Lee A. Dyer, Committee Member

Christopher R. Feldman, Committee Member

Stephen B. Vander Wall, Committee Member

Elizabeth A. Leger, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

August, 2015

Abstract

The purpose of this dissertation was to understand how changes in short-term ecological factors have lead to long-term evolutionary consequences, with a particular focus on the diversification of herbivorous insects and the plants they feed on. I begin with a general discussion of current theory on how changes to a species' niche, or general way of life, can shape evolutionary processes across a broad diversity of organisms, including adaptation, selection, diversification, and hybridization. Additionally, I discuss two important and unresolved questions specifically regarding the diversification of herbivorous insects and discuss how the subsequent three dissertation chapters attempt to address these areas of investigation. In Chapter 1, I investigate the evolutionary history of Neotropical moths in the hyperdiverse genus *Eois* (Geometridae), with an emphasis on documenting the roles of host conservatism, geography, and elevation in promoting diversification. In Chapter 2, I examine the evolution of secondary defense chemistry in one clade of *Piper* plants (*Radula*), the predominant genus that *Eois* caterpillars feed on. Finally in chapter 3, I elucidate patterns of novel host using Californian butterflies and exotic plants as a model system in an effort to determine the characteristics of herbivores that promote host shifting, which is typically the predominant mechanism by which diversification is thought to occur in herbivorous insect lineages.

Table of Contents

List of Tables

List of Figures

Introduction ~ Shifting ecological niches and their evolutionary consequences

Joshua P. Jahner

Program in Ecology, Evolution, and Conservation Biology, Department of Biology, University of Nevada, Reno, NV, 89557

Evolutionary consequences of shifting ecological niches

Within *On the Origin of Species*, Darwin (1859) highlighted the role of ecologically relevant differences in promoting natural selection and subsequent diversification, which has been a commonplace topic of evolutionary investigation ever since. Divergent ecological selection is now known to promote rapid ecological speciation over short time spans (Nosil 2012), resulting in the genesis of ecologically differentiated forms from a common ancestral form. One hallmark example is from stickleback fishes, where divergent selection on individuals occupying limnetic and benthic ecological niches has resulted in replicated divergence in multiple postglacial lakes in North America (Rundle et al. 2000). Similarly, selection acting on *Timema* walking sticks feeding on different host plants has led to two morphological phenotypes; each morph is best camouflaged on its respective host plant due to differences in host morphology (Gompert et al. 2014; Soria-Carrasco et al. 2014). In addition to adaptation to the niche itself, ecological selection can also act on a species' phenotype depending on the species composition found within a given community, as seen in North American velvet ants. These wingless wasps form one of the world's largest mimicry complexes, consisting of hundreds of species from multiple genera that fit into one of eight distinct color patterns (i.e., mimicry rings) (Wilson et al. in press). The mimetic fidelity of velvet ant species within a mimicry ring (i.e., how well species are adapted to the ring's mean phenotype) is hypothesized to largely depend on the diversity of other mimicry rings within a community (Wilson et al. 2013). More specifically, ecological selection is

thought to promote imperfect mimicry in highly diverse communities because imperfect mimics gain a benefit from matching multiple phenotypes, while higher mimetic fidelity is favored in less morphologically diverse velvet ant communities (Wilson et al. 2013). While ecological divergent selection is now known to be a strong driver of speciation across a number of lineages (Nosil 2012), it is important to note that other evolutionary processes promoting diversification are still influential in many of these systems, such as geographic effects. Future studies investigating the interactions between isolation by distance and isolation by environment (e.g., Wang et al. 2013) are still needed to fully understand the role that changes in a species niche play in diversification.

Ecological changes in niches are also thought to play a large role in driving diversification across deeper evolutionary scales. For adaptive radiations, which are characterized by the rapid multiplication of ecologically differentiated forms, ecological opportunity is thought to promote diversification by allowing a species to utilize a novel adaptive zone (i.e., a set of available niches that a species can utilize) (Simpson 1953; Schluter 2000; Losos 2010; Yoder et al. 2010). Simpson (1953) originally highlighted three routes to ecological opportunity, each associated with a change in a species' niche: colonization of a novel habitat, escape from antagonists, and the utilization of a novel resource type. Following each scenario, speciation is thought to be facilitated via divergent selection acting on individuals occupying different niches (Schluter 2000). Much work has been committed to understanding whether divergent ecological selection within an adaptive radiation results in a pattern where closely related organisms share similar niches or morphologies (i.e., phylogenetic signal; Blomberg and Garland 2002) or whether these traits evolve convergently, resulting in the pattern of distant relatives

sharing more similar traits. Numerous examples of convergent evolution have been documented across a variety of lineages, including *Anolis* lizards (Mahler et al. 2013), cichlid fishes (Muschick et al. 2012), and diurnal velvet ants (Wilson et al. in press), with strong ecological selection acting on lineages occupying alternate niches invoked as an active process driving diversification.

Shifting ecological factors have also been implicated in changing rates of gene flow among closely related species. Altered hybridization rates can occur in response to biotic changes to an environment, often in the form of shifting species assemblages. For instance, the rates and outcomes of hybridization among multiple western North American sucker species (*Catostomus* sp.) largely depends on which species occur within a given river drainage (Mandeville et al. 2015). More specifically, the native flannelmouth sucker (*C. latipinnis*) and the introduced white sucker (*C. commersoni*) hybridize extensively in the Little Sandy River; however, hybridization between these two species is dramatically reduced in the Big Sandy River where the longnose sucker (*C. catostomus*) has also been introduced (Mandeville et al. 2015). Numerous other examples of introduced species hybridizing with natives have been documented, often resulting in the introgression of introduced alleles into native species. Introgression of alleles can lead not only to morphological changes and reduced fitness in some cases (Muhlfield et al. 2009), but also has sparked a political and philosophical debate regarding the conservation of native species that contain exotic alleles (Fitzpatrick et al. 2010). Changes to a species' abiotic environment are also known to influence the rates and outcomes of hybridization among closely related species. For example, changes in species ranges in response to climate change can result in the formation of a new hybrid

zone, as has been reported in North American flying squirrels (Garroway et al. 2010). Furthermore, existing hybrid zones can shift geographically to match changing climates, in some cases resulting in differential introgression of alleles between the species (Taylor et al. 2014). Fluctuating weather patterns can also affect hybridization rates indirectly by altering the local abundances of parental species (Jahner et al. 2012). Finally, it is worth noting that both biotic and abiotic environmental changes can increase hybridization rates so dramatically that reproductive barriers between diverges species completely break down, potentially resulting in speciation reversal (Seehausen et al. 1997, 2008; Taylor et al. 2006).

What are the relative roles of biotic and abiotic changes in shaping long-term evolutionary processes?

It is clear from the examples above that both biotic and abiotic changes to a species' niche that occur during an ecologically relevant time scale can ultimately lead to profound evolutionary consequences. However, examining biotic and abiotic forces separately for a system might be misguided, as these forces are often not mutually exclusive from one another. While abiotic changes might directly promote evolution in a species, there often is an associated indirect effect on other biotic factors within the environment. This interplay between abiotic and biotic forces was readily apparent in the evolution of *Eois*, a hyperdiverse genus of Neotropical moths (Geometridae) that are host specialists on chemically well-defended plants in the genus *Piper* (Piperaceae). The evolution of *Eois* is characterized by repeated radiations associated with the colonization

of a novel host plant species, geographic area, or elevational band (see Chapter 1). Furthermore, trait differences found among *Eois* sister species are typically differences in host plant species, though some sister species also vary in geography or elevation (see Chapter 1). In sum, the results from *Eois* suggest that both biotic and abiotic changes to species niches during the evolution of the genus have led to diversification and that these niche changes are also associated with more recent divergence among sister species. For herbivorous insects, a widespread comparison of studies that find a strong role for host associated differentiation with studies examining lineages where host associated differentiation is not thought to play a role (e.g., in grass-feeding Lepidoptera; Jahner et al. 2015) could lead to better inferences about the relative roles of biotic and abiotic processes in shaping diversification and other evolutionary processes.

What factors explain the variation in changeability found across species and lineages?

There is widespread variation in the propensity to change ecological niches found both within and among lineages of herbivorous insects. In these diverse lineages, the colonization of a novel host plant has long been characterized as a shift to a novel ecological niche, allowing for subsequent diversification via divergent ecological selection (Ehrlich and Raven 1964). An understanding of the factors that promote host shifts is critical for a thorough understanding of evolutionary processes occurring at short time scales and macroevolutionary time scales. Interactions between native insects and exotic host plants represent a good system to understand variation in the ability to utilize a novel niche. For a set of 70 Californian butterflies that are known to utilize exotic host

plants, more generalist butterflies (in terms of native diet breadth) were more likely to utilize more exotic hosts (Chapter 3 / Jahner et al. 2011). This result is not surprising, as one might expect that species with a more constrained host range would be less likely to utilize a novel host. Additionally, more widespread Californian butterflies were more likely to utilize novel hosts, and this effect was stronger than the effect of native diet breadth (Chapter 3 / Jahner et al. 2011).

At deeper evolutionary scales, it is also possible to examine the propensity for lineages to change niches by estimating the amount of phylogenetic signal found across a phylogeny for a given trait. Phylogenetic signal is a pattern where closely related taxa typically share more similar traits than distant relatives (Blomberg and Garland 2002). For herbivorous insects, it has long been documented that lineages typically consume closely related plants with a similar chemical profile (Ehrlich and Raven, 1964), a pattern often referred to as host conservatism and also consistent with strong phylogenetic signal. For caterpillars in the genus *Eois* (described above), I found strong phylogenetic signal in host use, with lineages typically consuming host plants from within the same clade of *Piper*, the primary host genus for *Eois* (Chapter 1). Additionally, geography and elevation also displayed similarly strong phylogenetic signal across the phylogeny (Chapter 1). These results suggest that niche shifts associated with host plants, elevational bands, or novel geographic regions were relatively rare throughout the evolutionary history of the genus, but also that these shifts were often followed by diversification.

Similarly, most of the variation in secondary defense chemistry for host plants is thought to be partitioned predominantly among clades (i.e., phylogenetic signal), as opposed to within clades, though the evolution of secondary chemistry has been

examined in few lineages (Johnson et al. 2014). However, the evolution of secondary chemistry across plants in the *Piper* clade *Radula* was characterized by high lability in traits (Chapter 2). These results provide an interesting dichotomy to the high degree of phylogenetic signal found for *Eois*, suggesting that rapid diversification in highly diverse lineages can be associated with species with either highly conserved or highly changeable niches. It is possible that these patterns are not entirely independent of one another, as evolutionary changes in the host plant chemistry could cascade up into higher trophic levels to affect the evolution of host use in the *Eois* herbivores (Forister and Feldman 2011). For instance, a lack of phylogenetic signal in host plant chemistry among closely related plants might promote host specialization, especially if closely related plants are more likely to contain a chemical profile that is intolerable to an insect herbivore. This dynamic would contribute to structuring the global pattern of high specialization in diet breadth found in herbivorous insects (Forister et al. 2015) How this interplay between lability in host chemistry might affect host conservatism and patterns of specialization across local and global scales remains an open question, and should be investigated in a variety of other plant-insect systems.

References

Blomberg SP, Garland Jr. T (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, **15**, 899-910.

Darwin CR (1859) *On the origin of species by means of natural selection*. John Murray, London, UK.

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

Fitzpatrick BM, Johnson JR, Kump DK, Smith JJ, Voss SR, Shaffer HB (2010) Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences USA*, **107**, 3606-3610.

Forister ML, Feldman CR (2011) Phylogenetic cascades and the origins of tropical diversity. *Biotropica*, **43**, 270-278.

Forister ML, Novotny V, Panorska AK, Baje L, Basset Y, Butterill PT, Cizek L, Coley PD, Dem F, Diniz IR, Drozd P, Fox M, Glassmire AE, Hazen R, Hrcek J, Jahner JP, Kaman O, Kozubowski TJ, Kursar TA, Lewis OT, Lill J, Marquis RJ, Miller SE, Morais HC, Murakami M, Nickel H, Pardikes N, Ricklefs RE, Singer MS, Smilanich AM, Stireman JO, Villamarín-Cortez S, Vodka S, Volf M, Wagner DL, Walla T, Weiblen GD, Dyer LA (2015) The global distribution of diet breadth in insect herbivores. *Proceedings of the National Academy of Sciences USA*, **112**, 442-447.

Garroway CJ, Bowman J, Cascaden TJ, Holloway GL, Mahan CG, Malcolm JR, Steele MA, Turner G, Wilson PJ (2010) Climate change induced hybridization in flying squirrels. *Global Change Biology*, **16**, 113-121.

Gompert Z, Comeault AA, Farkas TE, Feder JL, Parchman TL, Buerkle CA, Nosil P (2014) Experimental evidence for ecological selection on genome variation in the wild. *Ecology Letters*, **17**, 369-379.

Jahner JP, Bonilla MM, Badik KJ, Shapiro AM, Forister ML (2011) Use of exotic hosts by Lepidoptera: widespread species colonize more novel hosts. *Evolution*, **65**, 2719-2724.

Jahner JP, Shapiro AM, Forister ML (2012) Drivers of hybridization in a 66-generation record of *Colias* butterflies. *Evolution*, **66**, 818-830.

Jahner JP, Forister ML, Nice CC, Fordyce JA, Wilson JS, Murphy DD, Marion ZH, Shapiro AM (2015) Regional population differentiation in the morphologically diverse, elevationally widespread Nearctic skipper *Polites sabuleti*. *Journal of Biogeography*, **42**, 1787-1799.

Johnson MTJ, Ives AR, Ahern J, Salminen JP (2014) Macroevolution of plant defenses against herbivores in the evening primroses. *New Phytologist*, **203**, 267-279.

Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, **175**, 623-639.

Mahler DL, Ingram T, Revell LJ, Losos JB (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, **341**, 292-295.

Mandeville EG, Parchman TL, McDonald DB, Buerkle CA (2015) Highly variable reproductive isolation among pairs of *Catostomus* species. *Molecular Ecology*, **24**, 1856-1872.

Muhlfeld CC, Kalinowski ST, McMahon TE, Taper ML, Painter S, Leary RF, Allendorf FW (2009) Hybridization rapidly reduces fitness of a native trout in the wild. *Biology Letters*, **5**, 328-331.

Muschick M, Indermaur A, Salzberger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Current Biology*, **22**, 2362-2368.

Nosil P (2012) *Ecological speciation*. Oxford University Press, New York, NY.

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 206-308.

Schluter D (2000) *The ecology of adaptive radiation*. Oxford University Press, New York, NY.

Seehausen O, van Alphen JJM, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, **277**, 1808-1811.

Seehausen O, Takimoto G, Roy D, Jokela J (2008) Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*, **17**, 30-44.

Simpson GG (1953) *The major features of evolution*. Columbia University Press, New York, NY.

Soria-Carrasco V, Gompert Z, Comeault AA, Farkas TE, Parchman TL, Johnston JS, Buerkle CA, Feder JL, Bast J, Schwander T, Egan SP, Crespi BJ, Nosil P (2014) Stick insect genomes reveal natural selection's role in parallel speciation. *Science*, **344**, 738-742.

Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, **15**, 343-355.

Taylor SA, White TA, Hochachka WM, Ferretti V, Curry RL, Lovette I (2014) Climate-mediated movement of an avian hybrid zone. *Current Biology*, **24**, 671-676.

Wang IJ, Glor RE, Losos JB (2013) Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters*, **16**, 175-182.

Wilson JS, Jahner JP, Williams KA, Forister ML (2013) Ecological and evolutionary processes drive the origin and maintenance of imperfect mimicry. *PLoS ONE*, **8**, e61610.

Wilson JS, Jahner JP, Forister ML, Sheehan ES, Williams KA, Pitts JP (in press) North American velvet ants form one of the world's largest known Müllerian mimicry complexes. *Current Biology*

Yoder JB, Clancey E, des Roches S, Eastman JM, Gentry L, Godsoe W, Hagey TJ, Jochimsen D, Oswald BP Robertson J, Sarver BAJ, Schenk JJ, Spear SF, Harmon LJ (2010) Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, **23**, 1581-1596.

Chapter 1 ~ Host conservatism, geography, and elevation in the evolution of a Neotropical moth radiation

Joshua P. Jahner[1,2], Matthew L. Forister[1,2], Angela M. Smilanich[1,2], Thomas L. Parchman[1,2], Joseph S. Wilson[3], Eric J. Tepe[4], Lora A. Richards[1,2], Mario Alberto Quijano-Abril[5], Christopher S. Jeffrey[1,6], Andrea E. Glassmire[1,2], Lee A. Dyer[1,2,7]

[1]*Program in Ecology, Evolution, and Conservation Biology, University of Nevada, Reno, NV, 89557;* [2]*Department of Biology, University of Nevada, Reno, NV, 89557;* [3]*Department of Biology, Utah State University, Tooele, UT 84074;* [4]*Department of Biological Sciences, University of Cincinnati, Cincinnati, OH 45221;* [5]*Grupo de Estudios Florísticos, Herbario Universidad Católica de Oriente, Rionegro, Antioquia, Colombia;* [6]*Department of Chemistry, University of Nevada, Reno, NV 89557;* [7]*Seccion Invertebrados, Museo Ecuatoriano de Ciencias Naturales, Quito, Ecuador;*

Keywords: Andes Mountains; Central American Seaway; ecological opportunity; *Eois*; genotyping-by-sequencing; *Piper*

Abstract

Adaptive radiations are thought to be stimulated by the colonization of a novel adaptive zone, such as an unoccupied habitat or unutilized resource type. For most herbivorous insects, the predominant mechanism of diversification is assumed to be a host shift onto a novel host plant lineage. However, other drivers of diversification also likely play an important role in shaping the evolutionary history of diverse lineages, especially in groups residing in regions with complex geological histories. In this study, we evaluated the contributions of host plant clade, geography, and elevation in shaping diversification in *Eois* (Lepidoptera: Geometridae), a hyper-diverse genus of Neotropical moths. Evolutionary relationships among 109 individuals were reconstructed using one mitochondrial gene (cytrochrome *c* oxidase subunit I) and two nuclear genes (elongation factor 1-alpha; wingless). As a complementary approach, an additional phylogeny was constructed from a genotyping-by-sequencing dataset encompassing 8,556 single nucleotide polymorphisms from 152 individuals. Both sequencing methodologies yielded similar topologies and strongly clustered individuals by host clade, geography, and elevation. While diversification of basal lineages often coincided with host clade shifts, the evolutionary history of *Eois* coincided strongly with major geological events, such as the rise of the Andes Mountains and the closure of the Central American Seaway. Overall, patterns of diversification in *Eois* are consistent with the perspective that shifts across multiple adaptive zones have driven diversification in this hyper-diverse lineage.

INTRODUCTION

Adaptive radiations are some of the most notable examples demonstrating the potential for ecological processes to shape long-term evolutionary dynamics (Simpson 1953; Schluter 2000). These systems are characterized by the rapid diversification of ecologically differentiated species, occasionally resulting in convergent evolution of similar forms in distant lineages (Muschick et al. 2012; Mahler et al. 2013; Wilson et al. 2015). Adaptive radiations can be stimulated by ecological opportunity associated with the invasion of a novel adaptive zone (Simpson 1953; Losos 2010; Yoder et al. 2010), often via dispersal into a previously unoccupied habitat or island (e.g. Lerner et al. 2011; Haines et al. 2014). Additionally, ecological opportunity can arise from the development of a key evolutionary innovation that allows for the colonization of a previously unutilized resource type (e.g. Hodges and Arnold 1995; Martin and Wainwright 2011). For instance, North American crossbills (*Loxia*) have morphologically specialized bills that are adept at opening the cones of conifers, and specialization of bill types on different conifer cone morphologies has led to rapid diversification (Parchman et al. 2006). While the role of ecological opportunity in driving diversification has been identified for many adaptive radiations, a comparison of multiple adaptive zones (i.e., a set of available niches that a species can utilize; Simpson 1953) has only been attempted in a few of the most well-studied groups (e.g. Givnish et al. 2009; Mahler et al. 2013). In this study, we tackle a hyper-diverse tropical radiation ripe for such investigation.

For herbivorous insects, we tend to assume that host plant lineages are the most important adaptive zone promoting diversification (e.g. Mitter et al. 1988; Fordyce 2010), which is readily apparent in hypotheses explaining the extreme diversity of phytophagous

insects and the plants they feed upon (Ehrlich and Raven 1964; reviewed by Janz 2011). In fact, the dominant paradigm is called the escape and radiate hypothesis, which posits that insect lineages often diversify on novel lineages of host plants after developing physiological adaptations to mitigate plant defenses (e.g. Wheat et al. 2007; Edger et al. 2015), suggesting a strong role for adaptive processes. However, one of the central patterns that led to the formulation of the escape and radiate hypothesis is that groups of closely related herbivores typically eat similar host plants (e.g. Janz et al. 2001; Stone et al. 2009; Wilson et al. 2012; Nylin et al. 2014; reviewed by Winkler & Mitter 2009), a pattern referred to as host conservatism. While host conservatism could be indicative of widespread ecological adaptation on related hosts, it is also an expected outcome of non-adaptive allopatric differentiation combined with strong phylogenetic niche conservatism (Wiens 2004). Despite the widespread evidence for host conservatism in herbivorous insect radiations, few studies have explicitly examined the relative importance of host conservatism and geographic differentiation for diversification (e.g. Becerra 1999; Condamine et al. 2012).

The Neotropical moth genus *Eois* (Lepidoptera: Geometridae) represents an ideal system for delineating the effects of host conservatism and geography for diversification of herbivorous insects. More than 250 species of *Eois* have been formally described, but the true diversity of the genus is estimated at more than 1,000 species in the Neotropics alone (Brehm et al. 2011). *Eois* caterpillars are highly specialized feeders, typically feeding on only one or two species (Connahs et al. 2009) from the genus *Piper* (Piperaceae), though associations have also been documented on other plant genera (Strutzenberger et al. 2010). Previous molecular investigations of *Eois* have reported host

plant conservatism, with lineages of related caterpillars consuming similar hosts (Strutzenberger et al. 2010; Wilson et al. 2012). Differences in elevational distributions, especially associated with the Andes Mountains, are also thought to promote diversification in *Eois* (Strutzenberger & Fiedler 2011), resulting in elevationally stratified communities of *Eois* (Rodríguez-Castañeda et al. 2010). Overall, two of the most important drivers of diversification in *Eois* that have been documented thus far are Andean uplift and the diversification of *Piper*, both of which occurred during the period with highest rates of *Eois* diversification (Strutzenberger & Fiedler 2011).

While many patterns of host-associated diversification have been documented in previous molecular analyses of *Eois* (e.g. Strutzenberger et al. 2010; Wilson et al. 2012), these studies largely included specimens from only three collection locations (Southern Ecuador, Central Ecuador, and Costa Rica) even though *Eois* can be found in the Neotropics from southern Mexico to northern Argentina (Brehm et al. 2011). This distributional range spans regions that have been dramatically affected by complex geological change during the past 50 million years, including the rise of the Andes Mountains (Hoorn et al. 2010) and the closure of the Central American Seaway that separated North and South America (Montes et al. 2015). These geological features have been implicated in shaping the phylogeographic histories of numerous Central and South American organisms from a variety of taxonomic groups (Hoorn et al. 2010; Turchetto-Zolet et al. 2013; Bagley and Johnson 2014) and have likely also impacted the evolutionary history of *Eois*. In this study we examine the phylogenetic relationships of *Eois* from across the Neotropics to investigate the roles of host conservatism, geography, and elevation in structuring patterns of diversification in *Eois* across the Neotropics.

Specifically, we test the predictions that 1) diversification is typically preceded by shifts in host plant clade, geography, or elevation; 2) host clade is the strongest force promoting differentiation, especially for ancestral nodes; and 3) recent diversification predominantly has occurred following shifts in geography or elevation. While our focus in this study is a broad-scale analysis of factors that could be involved in diversification (*sensu* Ehrlich and Raven 1964), we also consider the possibility that the same factors could be important at a small scale for the differentiation of sister species.

METHODS

Specimen collection

Eois caterpillars and moths were collected from sites in eight Central and South American countries or dependencies (Fig. 1A; Table S1), including individuals from the two sites previously examined by Wilson et al. (2012): Yanayacu Biological Station in Ecuador and La Selva Biological Station in Costa Rica. Sampling sites were selected to maximize the geographic, elevational, and host plant diversity of *Eois*. Specimens were collected via one of three methods: plot-based collecting, general collecting, or light trapping. Plot-based collecting involves searching for caterpillars on all *Piper* leaves in 10m diameter plots for one hour (see Rodríguez-Castañeda et al. 2010 for additional details) and is typically utilized at permanent sites. This method standardizes sampling effort when comparing ecological data across sites (e.g., Dyer et al. 2007; Forister et al. 2015). In contrast, general collection involves walking along transects and searching all encountered *Piper* plants for caterpillars. Caterpillars were either immediately preserved in ethanol or were reared to adult moths using standard protocols (Gentry and Dyer 2002)

to collect emerging parasitoid flies and wasps and for taxonomic identification. For each individual caterpillar, host plant specimens were also collected for identification and assignment to one of the eight Neotropical *Piper* host clades (roughly equivalent to a subgenus; Jaramillo et al. 2008; Fig. 1B). We were unable to assign host plant associations for adult moths collected via light trapping. In this study, we examined host conservatism at the clade level instead of the species level (e.g. Wilson et al. 2012) because we do not have genetic data for every host species and we were focused on broad scale patterns of evolution associated with the utilization of a novel host lineage (*sensu* Ehrlich and Raven 1964), rather than the effects of utilization of closely-related species.

Sanger sequencing

Eois DNA was extracted using Qiagen DNeasy Blood and Tissue kits (Qiagen Inc., Germantown, MD) and quantified using spectrophotometry. One mitochondrial gene (cytochrome *c* oxidase subunit I, COI) and two nuclear genes (elongation factor 1-alpha, EF1-α; wingless, WG) were amplified using PCR (see Table S2 for PCR primers and Table S3 for PCR protocols). Successfully amplified individuals were sequenced in both directions using an ABI 3730 DNA Analyzer (Applied Biosystems Inc., Foster City, CA) at the Nevada Genomics Center (Reno, NV, USA). Sequences were visually inspected and aligned in Sequencher 4.10.1 (Gene Codes Corp, Ann Arbor, MI) and submitted to Genbank.

Phylogenetic trees were first constructed for all three genes individually using Bayesian inference as implemented in MrBayes 3.2.3 (Ronquist et al. 2012). We were unable to exclusively use nominal taxonomy as a guide for the exclusion of duplicate

samples from phylogenetic analyses because cryptic genetic diversity has been reported in a number of tropical Lepidoptera lineages (e.g. Hebert et al. 2004; Burns et al. 2008). Furthermore, many of our samples were collected as caterpillars, which cannot be reliably identified to species at the larval stage (Wilson et al. 2012), and not reared to adulthood. Therefore, we only included individuals in phylogenetic reconstructions if they were different from one another in at least one of the following aspects: collection locality, host plant, morphology, nominal taxonomy, or genetic sequence. Models of evolution were selected by comparing AIC likelihood scores in jModeltest 2.1.5 (Darriba et al. 2012) for those models available in MrBayes. Based on the results from model selection, the GTR+I+G model of evolution was implemented for phylogenies constructed from nuclear genes (EF1-α and WG), while the HKY+I+G model of evolution was selected for the mitochondrial tree (COI). In addition to single gene trees, a phylogeny was constructed with all genes concatenated using a GTR+I+G model of evolution in an effort to fully resolve the topology. In order to root each tree, three outgroups were selected from a recent phylogenetic tree of the family Geometridae (Sihvonen et al. 2011): *Asthena albulata*, *Operophtera brumata*, and *Poecilasthena pulchraria*. All MrBayes analyses were run for two independent runs using 4 chains (3 heated and 1 cold) for 5,000,000 Markov chain Monte Carlo (MCMC) iterations. Chains were sampled every 2,500 MCMC iterations and a 25% burn-in was employed.

Genotyping by sequencing

As a complementary approach to traditional Sanger sequencing, we also constructed reduced-representation genomic libraries for Illumina sequencing using a genotyping-by-sequencing (GBS) approach (Parchman et al. 2012; Jahner et al. in review). DNA from the same extractions described above was cut at non-targeted sites throughout the genome using two restriction enzymes, EcoRI and MseI. Each individual *Eois* specimen was assigned a unique 8-10 base pair identifier to allow for highly multiplexed sequencing. Each DNA fragment was ligated to two adaptors: an EcoRI adaptor containing the Illumina adaptor, the individual's unique barcode, and bases matching the restriction enzyme cut site, and an MseI adaptor containing bases matching the cut site and the opposite Illumina adaptor. DNA libraries were amplified using PCR and fragments were size selected using a Blue Pippin quantitative electrophoresis unit (Sage Science, Beverly, MA). DNA was sequenced on two lanes of an Illumina HiSeq 2500 at the University of Texas Genomic Sequencing and Analysis Facility (Austin, TX).

Contaminant DNA (PhiX, *E. coli*), low quality fragments, and aberrant reads were filtered out of the Illumina dataset prior to assembly. In addition, we used a custom Perl script to correct single-base errors in barcode identifiers and to remove fragments containing Illumina adaptor fragments. A file was constructed with the unique barcode identifier inserted as the fastq identifier for each DNA fragment. We used SeqMan ngen software (DNASTAR Inc., Madison, WI) to construct an artificial reference genome from a subset of 25,000,000 DNA fragments, using a minimum match percentage of 93 and a gap penalty of 25. Contigs in the artificial reference genome were only included if they contained a minimum of 10 reads and 84-88 base pairs (over-assembled contigs were also removed). All other reads not included in the initial subset were subsequently aligned to

the artificial reference genome using the aln and samse algorithms in bwa (Burrows-Wheeler Aligner; Li and Durbin 2009).

We used samtools and bcftools (Li et al. 2009) to identify single nucleotide polymorphisms (SNPs) and to quantify the number of variable reads at each locus. For each individual, a genotype likelihood was calculated using bcftools (Li et al. 2009) for each genotype at every locus. These genotype likelihoods were stored in Variant Call Format (VCF) as composite genotype likelihoods, which account for uncertainty associated with sequencing errors and coverage, for use in subsequent analyses. SNPs were included in our analyses if only one alternative allele was present and if minor allele frequencies were greater than 5%. In addition, if fewer than 70% of *Eois* individuals contained at least one read at a locus, the SNP was excluded from analyses. The neighbor joining tree was constructed to characterize the relationships among *Eois* individuals included in GBS sequencing using the *bionj* function (Gascuel 1997) in the ape package (Paradis et al. 2004) in R v3.0.1 (R Core Team 2013), with the outgroup selected based on results from the Sanger sequencing results describe above. As a complementary approach, we also summarized genotypic variation across individuals using principal component analysis (PCA), implemented with the *prcomp* function in R.

Analyses of trait evolution

All analyses of trait evolution were conducted using both the concatenated and GBS trees. We first employed maximum likelihood to reconstruct ancestral character states for two discrete traits (Pagel 1994), country and host plant clade, to investigate patterns of diversification in *Eois*. An equal rates transition model was selected for

character reconstruction over a symmetric or all rates different model because the latter two models produced unreliable standard error estimates. Reconstruction was performed using the *ace* function of the ape package in R. In addition, we tested if host clade, country, and elevation (as a categorical trait; Fig. S1) were clustered non-randomly using three complementary analyses: parsimony score (PS; Slatkin & Maddison 1989), association index (AI; Wang et al. 2001), and maximum monophyletic clade size (MC; Parker et al. 2008). While PS and AI calculate a degree of clustering for an entire trait across a phylogeny, MC is a measure of clustering for each alternative trait state (e.g., Peru; *Radula*; mid elevation) (see Parker et al. 2008 for a detailed review of all three metrics). To test the null hypothesis that traits (PS and AI) and alternative trait states (MC) are randomly distributed across the phylogeny, observed values were compared to estimates from 1,000 tree-shuffling permutations. For all categorical analyses, elevation was categorized as low (<1,000m), medium (1,000-2,000m), and high (>2,000m) based on the elevational distribution of sampled individuals (Fig. S1). PS, AI, and MC values were calculated using the Bayesian Tip-Association Significance Testing (BaTS) software (Parker et al. 2008)

To test whether closely related *Eois* specimens share similar elevational distributions (as a continuous trait), we estimated two complementary metrics of phylogenetic signal, λ (Pagel 1999) and *K* (Blomberg *et al.* 2003) using the "phylosig" function of the phytools package (Revell 2012) in R. Phylogenetic signal refers to the tendency of closely related individuals to share similar trait values than more distant relatives (Blomberg & Garland 2002). Values of λ, a measure of the covariance among traits with respect to phylogenetic signal, range from zero (no signal) to one (strong

signal) (Pagel 1999). In contrast, *K* is calculated as the ratio of observed phylogenetic signal versus the expected signal under a model of Brownian motion, with values ranging from 0 to ∞ (values of $K < 1$ represent weak or absent signal) (Blomberg et al. 2003). We tested if estimates of λ were significantly different from 0 using a log-likelihood test comparing the observed estimate to an estimate of λ from a phylogenetic tree constrained to have λ=0. Significance for *K* was inferred by comparing the observed estimate to 1,000 tip-shuffling permutations.

Results

Sanger Sequencing

We sequenced a 617 bp COI fragment, a 679bp EF1-α fragment, and a 463bp WG fragment for 109 *Eois* individuals. The single gene trees were unable to fully resolve the phylogenetic history of *Eois*, especially along the backbone (Fig. 2). The COI gene tree provided the least phylogenetic resolution of the three genes, yielding almost no information regarding the backbone structure. In contrast, deeper resolution was found in the nuclear gene trees (EF1-α and WG), though each tree contained a large polytomy (Fig. 2). All three trees recovered a basal group composed of individuals feeding on the oldest Neotropical *Piper* clade, *Schilleria* (labeled "A" in Fig. 2). Some discordance was found across trees with regard to the placement of the "B" group, which was basal in the EF1-α tree and more apical in the WG tree, but these nodes did not receive strong support (Fig. 2). In general, the "C", "D", "E", "F", and "G" groups were composed of roughly the same individuals across the three gene trees, but the relative placement of these groups was unresolved due to the large polytomies.

Concatenation of the three genes resolved most of the nodes in the phylogenetic tree (Fig. S2) and clustered *Eois* individuals more strongly by country, elevation, and host clade than any of the single gene trees (Fig. 3). This qualitative assessment was confirmed by the estimates of AI and PS, as all three traits were significantly clustered on the phylogeny for both metrics (Table 1). For alternative trait states, the MC analysis revealed that all country, host clade, and elevation alternative states were significantly clustered on the concatenated tree, except for the *Isophyllon* clade ($N = 2$) (Table 2). Finally, estimates of phylogenetic signal for the elevational distribution of *Eois* were significantly greater than zero for both estimates ($\lambda = 0.998$; $P < 0.001$; $K = 0.148$; $P = 0.001$), though the K estimate was weak.

Genotyping by sequencing

Illumina sequencing yielded approximately 302 million sequences from the 152 *Eois* individuals. The artificial reference genome was constructed from the alignment of 7,529,213 reads during *de novo* assembly, forming 243,102 contigs. For neighbor-joining tree construction, we used 8,556 SNPs that had at least one read in >70% of individuals and minor allele frequencies greater than 5%. The GBS neighbor-joining tree (Fig. 4) yielded a similar topology to the concatenated tree (Fig. 3) and also strongly clustered individuals by country, geography, and host clade (Table 1). The major difference in topologies involved the placement of the small group labeled “C” in Fig. 2; “C” was placed sister to “D” in the concatenated tree (Fig. 3), while the GBS tree recovered “C” as sister to the “F” and “G” groups (Fig. 4).

All countries included in the MC analysis with at least three individuals were significantly clustered on the GBS tree, as well as all host clades except *Pothomorphe* (Table 2). However, only the low elevation category displayed significant clustering in the GBS dataset (Table 2). Estimates of elevational phylogenetic signal for the GBS tree ($\lambda = 0.908$; $P < 0.001$; $K = 0.093$; $P = 0.001$) were similar in magnitude to estimates from the concatenated tree. The first two PCs explained 18.59% and 5.68% of the variance in genotype likelihoods, respectively. The PCA was generally unable to delineate groups of individuals by country, host clade, or geography (Fig. 6), which is not surprising given the broad convergence in traits across the *Eois* phylogeny (Figs. 3 & 4). However, individuals from a few categories did cluster together strongly in PC space, including French Guiana, Panama, Peru, *Peltobryon*, *Schilleria*, and high elevation (Fig. 6).

Discussion

In this study, we reconstructed four phylogenies from traditional sequencing methods (Figs. 2 & 3), as well as a phylogeny using a GBS approach (Fig. 4), to characterize the drivers of adaptive radiation in a hyper-diverse genus of Neotropical moths. Similar topologies were recovered for both datasets (Figs. 3 & 4) and individuals were clustered together by three traits hypothesized to be associated with diversification in herbivorous insect lineages: host plant clade, geography, and elevation (Tables 1 & 2). Despite the relatively low number of SNPS ($N = 8,556$) we recovered due to the high sequence divergence found across the genus, the concordance in both topology and patterns of trait evolution suggests that GBS protocols might be an extremely useful alternative to more traditional phylogenetic sequencing methods (e.g. Wagner et al. 2013;

Ebel et al. 2015). Overall, the evolutionary history of *Eois* is characterized by shifts in host clade, geography, or elevation, followed by subsequent diversification (Figs. 3 & 4). The patterns of adaptive radiation in *Eois* are thus consistent with the escape and radiate hypothesis (Ehrlich and Raven 1964) and, in a broader sense, ecological opportunity promoting diversification across multiple adaptive zones.

The results from the phylogenetic reconstruction of host clade use in *Eois* conform to the expectations of an escape and radiate mechanism of diversification (Ehrlich & Raven 1964). Within most large lineages, host clade is highly conserved (Figs. 3 & 4) suggesting that major host clade shifts likely occurred early in the diversification of *Eois* and were followed by diversification of lineages feeding on the same host clade. These results support the hypothesis of Wilson et al. (2012), who posited that Neogene diversification of *Eois* was spurred by host shifts, largely coinciding with the diversification of *Piper*. In addition to finding support for host clade shifts driving Neogene diversification, a number of sister species pairs differed in host clade (Figs. 3 & 4), indicating that this mode of diversification is still an active process in some *Eois* lineages. This result is somewhat unexpected given the extreme host specificity of *Eois*, where each species typically consumes only one or two species of *Piper* (Connahs et al. 2009), because diet breadth is positively correlated with the propensity to utilize novel hosts in other lineages (Jahner et al. 2011). Finally, many of the most basal individuals on the trees consumed plants in *Schilleria* (Figs. 3 & 4), the most basal Neotropical *Piper* clade (Fig. 1), consistent with codiversification of *Eois* and their host plants.

The oldest *Eois* lineages are comprised mostly of individuals found in present day Panama (Figs. 3 & 4), which is surprising given that most of Panama was thought to be submerged under the Central American Seaway (CAS) that divided North and South America until the middle Miocene (~13-15 Ma) (Montes et al. 2015). This result suggests that either ancestral lineages of *Eois* dispersed into Panama following the closure of the CAS or that the closure of the CAS was dynamic, with areas of present day Panama remaining above sea level and acting as refugia until the final closure. Evidence supporting the latter scenario (complex closure of the CAS) has also been reported in a phylogenetic reconstruction of North and South American bees in the genus *Diadasia*, which found evidence for dispersal between the two continents approximately 15-20.5 Ma (Wilson et al. 2014). The geographic distribution among more apical *Eois* lineages does not suggest the CAS has been a recent barrier to dispersal, with repeated changes in geographical distribution between Central and South America found throughout the phylogenies (Figs. 3 & 4). Despite the high number of geographic shifts among lineages, country was significantly clustered for both datasets (Tables 1 & 2), suggesting that *Eois* lineages often diversify in relatively close geographic proximity, though geographic shifts are also associated with recent diversification (Fig. 5).

The distribution of elevational ranges across the phylogenies supports a low-elevation origin of *Eois* followed by the subsequent repeated colonization of higher elevations (Figs. 3 & 4). This pattern is readily apparent in the GBS tree, where the large majority of basal individuals occupied elevations lower than 1,000m (Fig. 4). Elevational clustering is less pronounced in the Sanger dataset, where a higher number of mid-elevation individuals were found interspersed in lineages of low-elevation individuals in

the basal portions of the tree (Fig. 3). Previous molecular clock analyses have found results suggesting that much of the diversification in *Eois* coincided with increased Andean uplift, particularly in the Neogene (Strutzenberger & Fiedler 2011), which is consistent with our results. It is important to note, however, that the genus *Piper* also rapidly diversified during the rise of the Andes (Wilson et al. 2012), so uplift likely triggered accelerated *Eois* differentiation both directly through the genesis of novel habitats and indirectly by promoting host plant diversification. For *Eois* caterpillars, low and high elevations can be considered different ecological niches because attack rates from natural enemies (e.g. predatory ants or parasitoid flies and wasps) are greater in low elevations (O'Donnell & Kumar 2006; Connahs et al. 2009; Rodríguez-Castañeda et al. 2011), resulting in stratification of *Eois* communities across elevational bands (Rodríguez-Castañeda et al. 2010). Thus, the colonization of a higher elevational band followed by diversification could have been driven by ecological opportunity (Losos 2010; Yoder et al. 2010), with high elevation species likely released from the strong enemy pressures characteristic of the lowlands. Overall, the patterns of elevational differentiation add to the growing list of studies documenting the role of elevational gradients in driving diversification of butterflies and moths around the world (e.g., Elias et al. 2009; Karl et al. 2009; Casner & Pyrcz 2010; Jahner et al. 2015)

While the majority of sister species pairs found in the GBS tree differed in at least one trait (host species, country, or elevation band), 15 of the pairs were invariant (Fig. 5). One explanation for the apparent lack of trait differentiation is that diversification between these pairs is still incipient, and differentiation has not yet occurred in the traits examined in this study. Another possibility is that our categories might be too coarse-

grained to distinguish the differences among closely related individuals. Alternatively, other traits that were not investigated in this study could also play a role in driving diversification in *Eois*. For instance, phytochemical variation across *Piper*, the primary host plant genus, is highly variable even between closely related species (Jahner et al. in prep). Thus, *Eois* may be diversifying along a phytochemical niche axis not quantified here, which would also support the perspective that diversification in this group is driven by ecological opportunity associated with multiple adaptive zones synergistically. It is important to note that while we have found strong evidence for shifts in traits promoting diversification, the strong phylogenetic clustering of *Eois* with shared traits (Tables 1 & 2) is also evidence for phylogenetic niche conservatism (Wiens 2004), which could explain differentiation in the sister pairs with invariant traits. It is highly likely that both niche conservatism and shifts in adaptive zones are important in driving more recent diversification in *Eois.*

In this study, we investigated the roles of host conservatism, geography, and elevation in driving the diversification of a hyper-diverse genus of specialized moths. Our results align with the expectations originally outlined by Ehrlich and Raven (1964), namely that rapid diversification of insect lineages feeding on similar hosts are often preceded by shifts to novel host lineages. However, our results also support an equal role for shifts in geographic and elevational distributions in driving diversification, with recent diversification between a number of sister species associated with a shift in host species, country, or elevational band (Fig. 5). These findings suggest that future studies examining diversification in herbivorous insects must account for geographic and

elevational patterns in addition to host use patterns, especially for lineages with distributions residing in regions with geologically complex histories.

Acknowledgements

This research was largely funded through a grant from the National Science Foundation (DEB-1145609). We thank Chris Feldman, Beth Leger, and Steve Vander Wall for promiding valuable comments that have improved the text.

References

Bagley JC, Johnson JB (2014) Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews*, **89**, 767-790.

Becerra JX (1999) Macroevolution of insect-plant associations: the relevance of host biogeography to host affiliation. *Proceedings of the National Academy of Sciences USA*, **96**, 12626-12631.

Blomberg SP, Garland Jr. T (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, **15**, 899-910.

Blomberg SP, Garland Jr. T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, **57**, 717-745.

Brehm G, Bodner F, Strutzenberger P, Hünefeld F, Fiedler K (2011) Neotropical *Eois* (Lepidoptera: Geometridae): checklist, biogeography, diversity, and description patterns. *Annals of the Entomological Society of America*, **104**, 1091-1107.

Burns JM, Janzen DH, Hajibabaei M, Hallwachs W, Hebert PDN (2008) DNA barcodes and cryptic species of skipper butterflies in the genus *Perichares* in Area de Conservación Guanacaste, Costa Rica. *Proceedings of the National Academy of Sciences USA*, **105**, 6350-6355.

Casner KL, Pyrcz TW (2010) Patterns and timing of diversification in a tropical montane butterfly genus, *Lymanopoda* (Nymphalidae, Satyrinae). *Ecography*, **33**, 251-259.

Condamine FL, Sperling FAH, Wahlberg N, Rasplus JY, Kergoat GJ (2012) What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecology Letters*, **15**, 267-277.

Connahs H, Rodríguez-Castañeda G, Walters T, Walla T, Dyer L (2009) Geographic variation in host-specificity and parasitoid pressure of an herbivore (Geometridae) associated with the tropical genus *Piper* (Piperaceae). *Journal of Insect Science*, **9**, 28.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Dyer LA, Singer MS, Lill JT *et al.* (2007) Host specificity of Lepidoptera in tropical and temperate forests. *Nature*, **448**, 696-699.

Ebel ER, DaCosta JM, Sorenson MD *et al.* (2015) Rapid diversification associated with ecological specialization in Neotropical *Adelpha* butterflies. *Molecular Ecology*, **24**, 2392-2405.

Edger PP, Heidel-Fischer HM, Bekaert M *et al.* (2015) The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences USA*

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

Elias M, Joron M, Willmott K *et al.* (2009) Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, **18**, 1716-1729.

Fordyce JA (2010) Host shifts and evolutionary radiations of butterflies. *Proceedings of the Royal Society B*, **277**, 3735-3743.

Forister ML, Novotny V, Panorska AK *et al.* (2015) The global distribution of diet breadth in insect herbivores. *Proceedings of the National Academy Sciences USA*, **112**, 442-447.

Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution*, **14**, 685-695.

Gentry GL, Dyer LA (2002) On the conditional nature of Neotropical caterpillar defenses against their natural enemies. *Ecology*, **83**, 3108-3119.

Givnish TJ, Millam KC, Mast AR *et al.* (2009) Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B*, **276**, 407-416.

Haines WP, Schmitz P, Rubinoff D (2014) Ancient diversification of *Hyposmocoma* moths in Hawaii. *Nature Communications*, **5**, 3502.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences USA*, **101**, 14812-14817.

Hodges SA, Arnold ML (1995) Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society B*, **262**, 343-348.

Hoorn C, Wesselingh FP, ter Steege H *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927-931.

Jahner JP, Bonilla MM, Badik KJ, Shapiro AM, Forister ML (2011) Use of exotic hosts by Lepidoptera: widespread species colonize more novel hosts. *Evolution*, **65**, 2719-2724.

Jahner JP, Forister ML, Nice CC *et al.* (2015) Regional population differentiation in the morphologically diverse, elevationally widespread Nearctic skipper, *Polites sabuleti*. *Journal of Biogeography* doi: 10.1111/jbi.12557

Jahner JP, Gibson D, Weitzman CL, Blomberg EJ, Sedinger JS, Parchman TL (In review) Lek geography drives fine-scale genetic structure of greater sage-grouse. *Molecular Ecology*

Jahner JP, Tepe EJ, Richards LA *et al.* (In prep) Phylogenetic signal and the evolution of secondary chemistry in *Radula* (Piperaceae): persistent trait correlations despite rampant trait lability.

Janz N (2011) Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annual Review of Ecology, Evolution and Systematics*, **42**, 71-89.

Janz N, Nyblom K, Nylin S (2001) Evolutionary dynamics of host-plant specialization: a case study of the tribe Nymphalini. *Evolution*, **55**, 783-796.

Jaramillo MA, Callejas R, Davidson C, Smith JF, Stevens AC, Tepe EJ (2008) A phylogeny of the tropical genus *Piper* using ITS and the chloroplast intron *psbJ-petA*. *Systematic Botany*, **33**, 647-660.

Karl I, Schmitt T, Fischer K (2009) Genetic differentiation between alpine and lowland populations of a butterfly is related to PGI enzyme genotype. *Ecography*, **32**, 488-496.

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC (2011) Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, **21**, 1838-1844.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760.

Li H, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.

Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, **175**, 623-639.

Mahler DL, Ingram T, Revell LJ, Losos JB (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, **341**, 292-295.

Martin CH, Wainwright PC (2011) Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of *Cyprinodon* pupfish. *Evolution*, **65**, 2197-2212.

Mitter C, Farrell B, Wiegmann B (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *The American Naturalist*, **132**, 107-128.

Montes C, Cardona A, Jaramillo C *et al.* (2015) Middle Miocene closure of the Central American Seaway. *Science*, **348**, 226-229.

Muschick M, Indermaur A, Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Current Biology*, **22**, 2362-2368.

Nylin S, Slove J, Janz N (2014) Host plant utilization, host range oscillations and diversification in nymphalid butterflies: a phylogenetic investigation. *Evolution*, **68**, 105-124.

O'Donnell S, Kumar A (2006) Microclimatic factors associated with elevational changes in army ant density in tropical montane forest. *Ecological Entomology*, **31**, 491-498.

Pagel M (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society London B*, **255**, 37-45.

Pagel M (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877-884.

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289-290.

Parchman TL, Benkman CW, Britch SC (2006) Patterns of genetic variation in the adaptive radiation of New World crossbills (Aves: *Loxia*). *Molecular Ecology*, **15**, 1873-1887.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle CA (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, **21**, 2991-3005.

Parker J, Rambaut A, Pybus OG (2008) Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. *Infection, Genetics and Evolution*, **8**, 239-246.

R Core Team (2013) R: language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 URL http://www.R-project.org/

Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217-223.

Rodríguez-Castañeda G, Dyer LA, Brehm G, Connahs H, Forkner RE, Walla TR (2010) Tropical forests are not flat: how mountains affect herbivore diversity. *Ecology Letters*, **13**, 1348-1357.

Rodríguez-Castañeda G, Forkner RE, Tepe EJ, Gentry GL, Dyer LA (2011) Weighing defensive and nutritive roles of ant mutualists across a tropical altitudinal gradient. *Biotropica*, **43**, 343-350.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539-542.

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

Sihvonen P, Mutanen M, Kaila L, Brehm G, Hausmann A, Staude HS (2011) Comprehensive molecular sampling yields a robust phylogeny for geometrid moths (Lepidoptera: Geometridae). *PLoS ONE*, **6**, e20356.

Simpson GG (1953) *The major features of evolution*. Columbia University Press, New York, NY.

Slatkin M, Maddison WP (1989) A cladistics measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603-613.

Stone GN, Hernandez-Lopez A, Nicholls JA *et al.* (2009) Extreme host plant conservatism during at least 20 million years of host plant pursuit by oak gallwasps. *Evolution*, **63**, 854-869.

Strutzenberger P, Fiedler K (2011) Temporal patterns of diversification in Andean *Eois*, a species-rich clade of moths (Lepidoptera, Geometridae). *Journal of Evolutionary Biology*, **24**, 919-925.

Strutzenberger P, Brehm G, Bodner F, Fiedler K (2010) Molecular phylogeny of *Eois* (Lepidoptera, Geometridae): evolution of wing patterns and host plant use in a species-rich group of Neotropical moths. *Zoologica Scripta*, **39**, 603-620.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary processes in South America. *Molecular Ecology*, **22**, 1193-1213.

Wagner CE, Keller I, Wittwer S *et al.* (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787-798.

Wang TH, Donaldson YK, Brettle RP, Bell JE, Simmonds P (2001) Identification of shared populations of Human Immunodeficiency Virus Type 1 infecting microglia and tissue macrophages outside the central nervous system. *Journal of Virology*, **75**, 11686-11699.

Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant-insect coevolutionary key innovation. *Proceedings of the National Academy of Sciences USA*, **104**, 20427-20431.

Wiens JJ (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, **58**, 193-197.

Wilson JS, Forister ML, Dyer LA *et al*. (2012) Host conservatism, host shifts and diversification across three trophic levels in two Neotropical forests. *Journal of Evolutionary Biology*, **25**, 532-546.

Wilson JS, Carril OM, Sipes SD (2014) Revisiting the Great American Biotic Interchange through analyses of amphitropical bees. *Ecography*, **37**, 791-796.

Wilson JS, Jahner JP, Forister ML, Sheehan ES, Williams KA, Pitts JP (2015) North American velvet ants form one of the world's largest known Müllerian mimicry complexes. *Current Biology*

Winkler IS, Mitter C (2009) The phylogenetic dimension of insect-plant interactions: a review of recent evidence. In: *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects* (ed. Tilmon KJ), pp. 203-215. University of California Press, Berkeley, CA.

Yoder JB, Clancey E, des Roches S *et al*. (2010) Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, **23**, 1581-1596.

Table 1. Observed Association Index (Wang et al. 2001) and Parsimony Score (Slatkin and Maddison 1989) measurements of trait clustering were compared to expected values from 1,000 tree-tip shuffling permutations to test the null hypothesis that traits were randomly distributed across phylogenies. Both analyses were conducted on country, host clade, and categorical elevation traits using both datasets (Sanger sequencing and GBS). Significant clustering was inferred if $P < 0.05$ (denoted by bold text).

Dataset	Trait	Analysis	Observed	Expected Mean (95% CI)	*P*
Sanger	Country	AI	2.63	9.37 (8.18 - 10.51)	**< 0.001**
		PS	33	57.47 (54 - 61)	**< 0.001**
	Clade	AI	2.72	8.93 (7.74 - 10.05)	**< 0.001**
		PS	29	53.79 (50 - 57)	**< 0.001**
	Elevation	AI	2.77	8.51 (7.25 - 9.80)	**< 0.001**
		PS	31	53.74 (49 - 58)	**< 0.001**
GBS	Country	AI	3.47	10.93 (9.50 - 12.22)	**< 0.001**
		PS	38	72.60 (68 - 77)	**< 0.001**
	Clade	AI	3.93	11.31 (9.92 - 12.59)	**< 0.001**
		PS	44	73.62 (70 -77)	**< 0.001**
	Elevation	AI	4.19	10.10 (8.47 - 11.58)	**< 0.001**
		PS	40	68.62 (63 - 74)	**< 0.001**

Table 2. Results from the maximum monophyletic clade (MC) size analyses (Parker et al. 2008), which tested the null hypothesis that alternative states for the traits country, host clade, and elevation were distributed randomly across the *Eois* phylogeny. A separate analysis was run for each trait, in which the observed MC was compared to a null MC from 1,000 tree randomizations for each alternative state. Significant clustering of traits (denoted by bold text) was inferred for alternative states that rejected the null hypothesis ($P < 0.05$). Both analyses were performed using phylogenies created from the Sanger sequencing (concatenated tree) and GBS datasets. The results from alternative states that were comprised of a single individual (Sanger/Country: Mexico; Sanger/Host: *Enckea*; GBS/Clade: *Enckea*, *Isophyllon*) are not reported.

Dataset	Trait	Alternative State	*N*	Observed MC	Null MC	*P*
Sanger	Country	Argentina	3	2	1.02	**0.020**
		Colombia	11	3	1.30	**0.015**
		Costa Rica	26	11	2.10	**< 0.001**
		Ecuador	49	10	3.33	**< 0.001**
		French Guiana	5	2	1.05	**0.049**
		Panama	8	4	1.14	**< 0.001**
		Peru	6	2	1.09	0.087
	Host Clade	*Isophyllon*	2	1	1.00	1
		Macrostachys	20	4	1.80	**0.006**
		Peltobryon	8	5	1.15	**< 0.001**
		Pothomorphe	6	4	1.07	**< 0.001**
		Radula	54	7	3.67	**0.015**
		Schilleria	5	5	1.05	**< 0.001**
	Elevation	Low	39	6	2.70	**0.008**
		Medium	42	6	2.85	**0.015**
		High	28	4	2.19	**0.027**
GBS	Country	Argentina	3	1	1.01	1
		Costa Rica	43	9	2.47	**< 0.001**
		Ecuador	69	8	3.64	**0.010**
		French Guiana	4	2	1.02	**0.021**
		Mexico	2	1	1.00	1
		Panama	16	4	1.41	**0.002**
		Peru	15	3	1.38	**0.026**
	Host Clade	*Hemipodium*	3	2	1.01	**0.010**
		Macrostachys	28	4	1.94	**0.010**
		Peltobryon	10	7	1.16	**< 0.001**
		Pothomorphe	9	2	1.12	0.120
		Radula	73	17	3.89	**< 0.001**
		Schilleria	9	3	1.13	**0.004**
	Elevation	Low	62	10	3.35	**0.002**
		Medium	57	5	3.01	0.073
		High	33	3	2.12	0.193

Figure Legends

Fig. 1. *Eois* caterpillars and moths were collected (A) from eight countries across the Neotropics, including (B) individuals feeding on all New World *Piper* clades (topology from Jaramillo et al. 2008). Colors match those used in Figs. 3, 4, & 6.

Fig. 2. The evolutionary relationships of 109 *Eois* were reconstructed using one mitochondrial gene (COI) and two nuclear genes (EF1-α; WG). Individuals not identified to species are labeled as either rare or common (depending on how many individuals were collected), followed by the host plant species name. Additional individual collection information can be found in Table S1. Node support values are listed for nodes with posterior probabilities >0.90. Nodes labeled with letters are discussed more thoroughly in the main text.

Fig. 3. The distribution of geographic, elevational, and host clade variation is displayed across the concatenated phylogeny (COI; EF1-α; WG). Individuals not identified to species are labeled as either rare or common (depending on how many individuals were collected), followed by the host plant species name. Additional individual collection information can be found in Table S1. Pie charts display the results of ancestral state reconstructions for (A) country and (E) host clade for each node. The bars in the center of the figure display current tip states for (B) country, (C) elevation, and (D) host clade. See Fig. S2 for node support values.

Fig. 4. The distribution of geographic, elevational, and host clade variation is displayed across the GBS neighbor-joining tree. Individuals not identified to species are labeled as either rare or common (depending on how many individuals were collected), followed by the host plant species name. Additional individual collection information can be found in Table S1. Pie charts display the results of ancestral state reconstructions for (A) country and (E) host clade for each node. The bars in the center of the figure display current tip states for (B) country, (C) elevation, and (D) host clade.

Fig. 5. (A) A histogram displays the number of sister-species pairs from the GBS dataset that vary by zero, one, two, or three traits (country, elevation, and host species) (total N = 44). (B) For each trait, the number of sister species pairs that share (grey) or differ (black) by country, elevation, and host species are displayed. Six pairs are not displayed for the host species comparison because the host species identity could not be determined.

Fig. 6. Results from the PCA of genotype likelihoods of 8,556 SNPs from the 152 individuals included in the GBS dataset. Values for each individual are colored by country, host clade, and elevation. PC1 and PC2 explained 18.59% and 5.68% of the variance in genotype likelihoods, respectively.

Figure 1

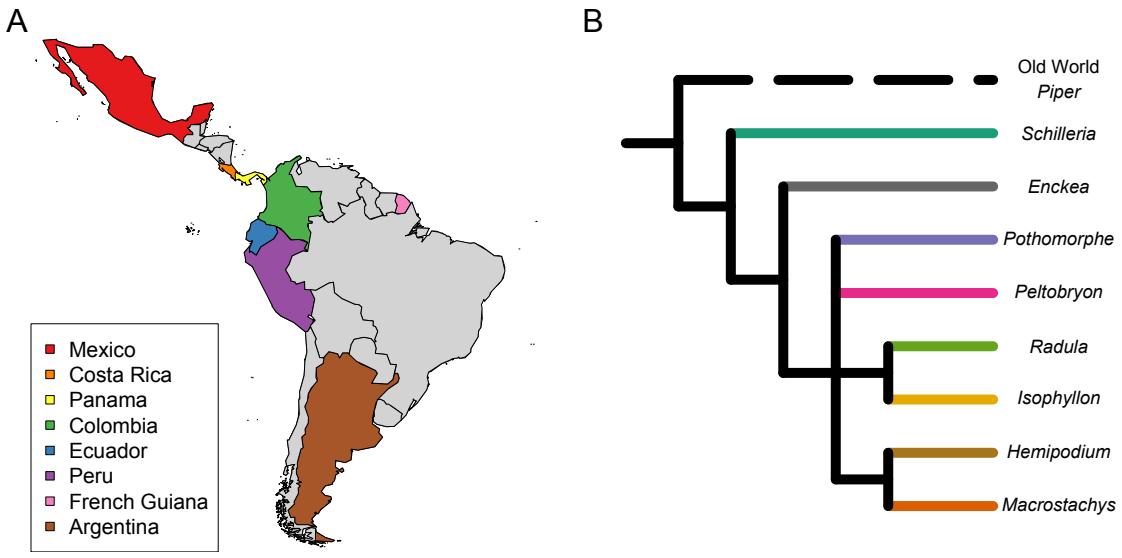

A
Mexico
Costa Rica
Panama
Colombia
Ecuador
Peru
French Guiana
Argentina
B
Old World
Piper
Schilleria
Enckea
Pothomorphe
Peltobryon
Radula
Isophyllon
Hemipodium
Macrostachys

Figure 2

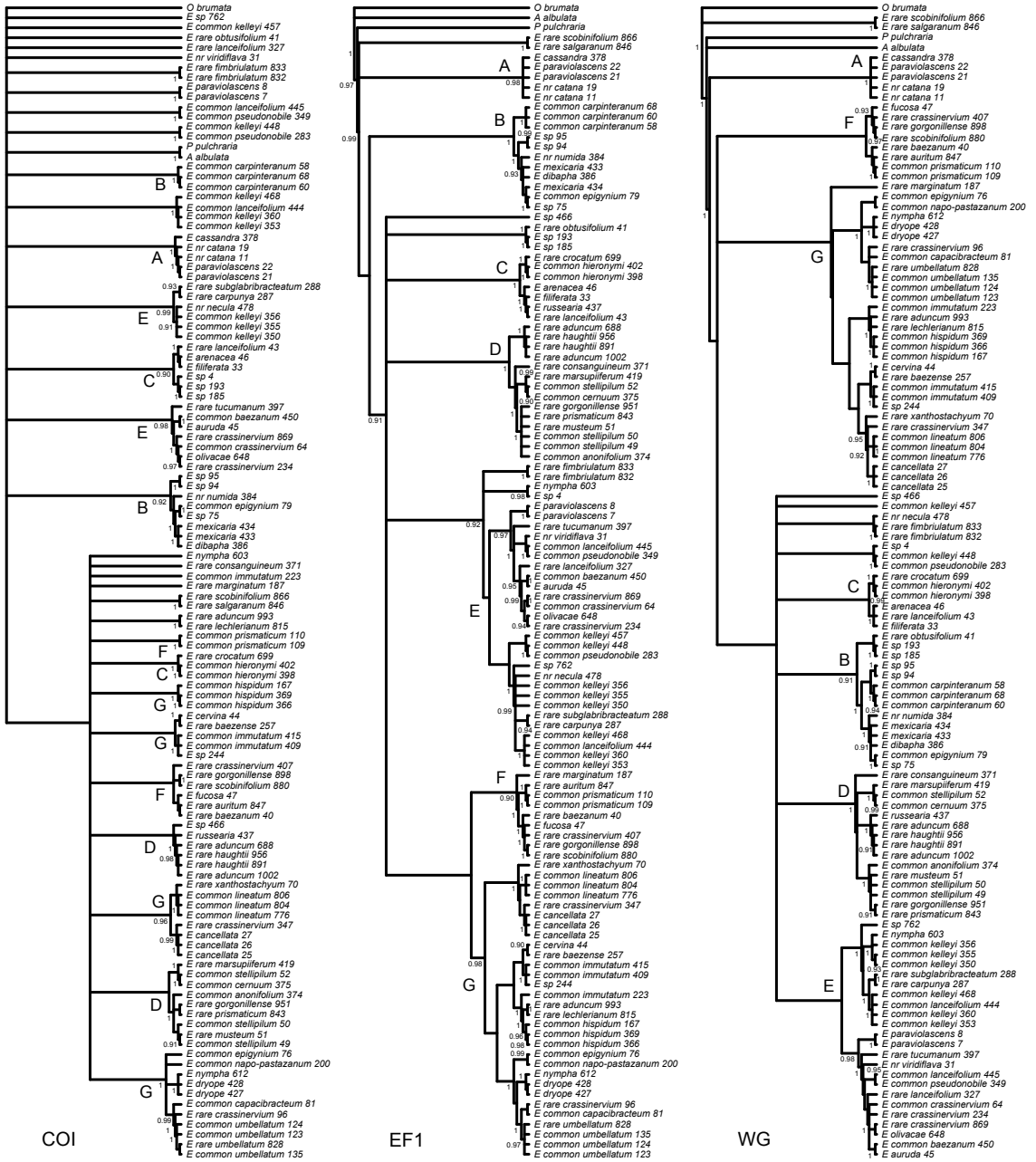


Figure 3

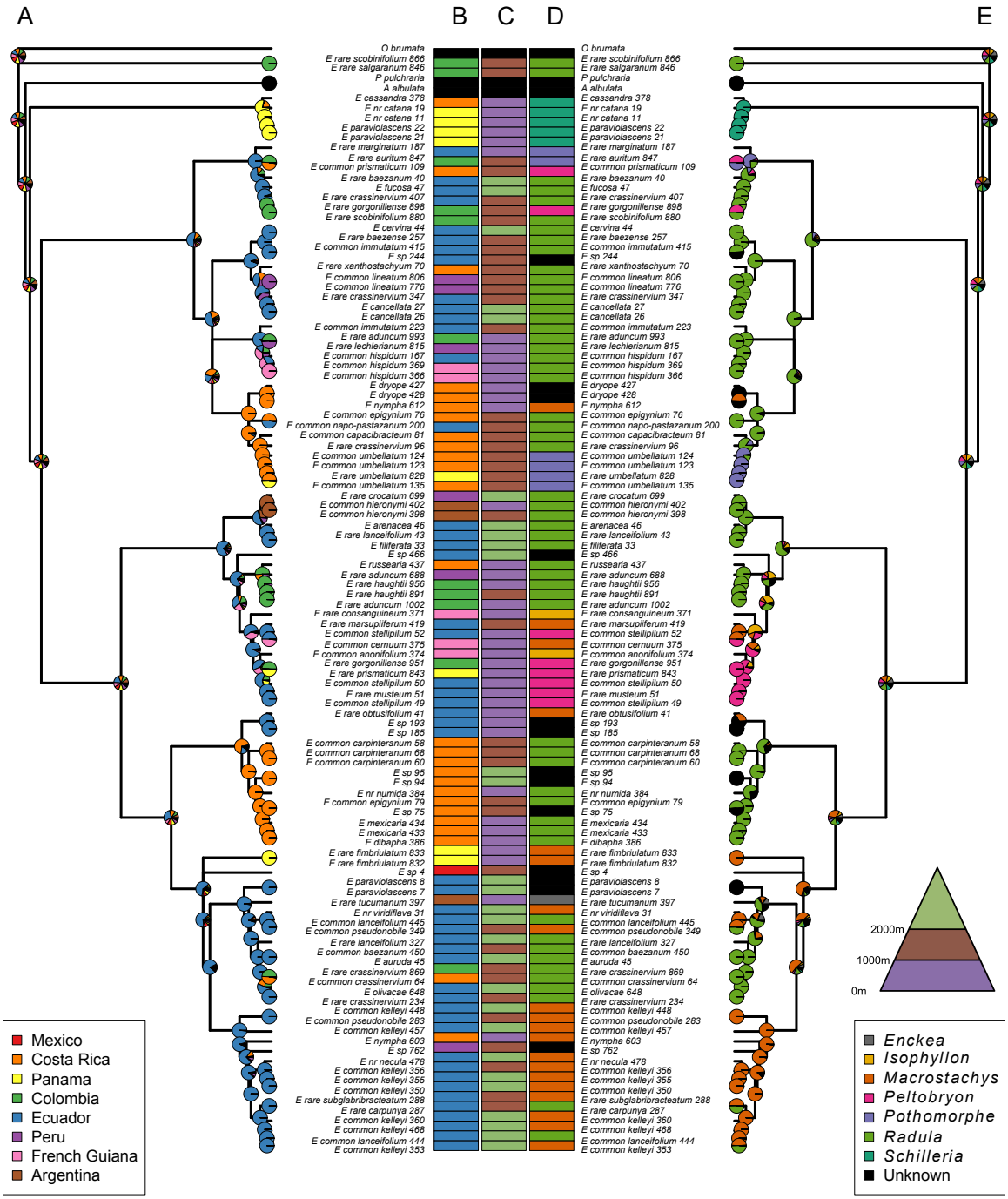

A
B
C
D
E
O brumata
E rare scobinifolium 866
E rare salgaranum 846
P pulchraria
A albulata
E cassandra 378
E nr catana 19
E nr catana 11
E paraviolascens 22
E paraviolascens 21
E rare marginatum 187
E rare auritum 847
E common prismaticum 109
E rare baezanum 40
E fucosa 47
E rare crassinervium 407
E rare gorgonillense 898
E rare scobinifolium 880
E cervina 44
E rare baezense 257
E common immutatum 415
E sp 244
E rare xanthostachyum 70
E common lineatum 806
E common lineatum 776
E rare crassinervium 347
E cancellata 27
E cancellata 26
E common immutatum 223
E rare aduncum 993
E rare lechlerianum 815
E common hispidum 167
E common hispidum 369
E common hispidum 366
E dryope 427
E dryope 428
E nympha 612
E common epigynium 76
E common napo-pastazanum 200
E common capacibracteum 81
E rare crassinervium 96
E common umbellatum 124
E common umbellatum 123
E rare umbellatum 828
E common umbellatum 135
E rare crocatum 699
E common hieronymi 402
E common hieronymi 398
E arenacea 46
E rare lanceifolium 43
E filiferata 33
E sp 466
E russearia 437
E rare aduncum 688
E rare haughtii 956
E rare haughtii 891
E rare aduncum 1002
E rare consanguineum 371
E rare marsupiiferum 419
E common stellipilum 52
E common cernuum 375
E common anonifolium 374
E rare gorgonillense 951
E rare prismaticum 843
E common stellipilum 50
E rare musteum 51
E common stellipilum 49
E rare obtusifolium 41
E sp 193
E sp 185
E common carpinteranum 58
E common carpinteranum 68
E common carpinteranum 60
E sp 95
E sp 94
E nr numida 384
E common epigynium 79
E sp 75
E mexicaria 434
E mexicaria 433
E dibapha 386
E rare fimbriulatum 833
E rare fimbriulatum 832
E sp 4
E paraviolascens 8
E paraviolascens 7
E rare tucumanum 397
E nr viridiflava 31
E common lanceifolium 445
E common pseudonobile 349
E rare lanceifolium 327
E common baezanum 450
E auruda 45
E rare crassinervium 869
E common crassinervium 64
E olivacae 648
E rare crassinervium 234
E common kelleyi 448
E common pseudonobile 283
E common kelleyi 457
E nympha 603
E sp 762
E nr necula 478
E common kelleyi 356
E common kelleyi 355
E common kelleyi 350
E rare subglabribracteatum 288
E rare carpunya 287
E common kelleyi 360
E common kelleyi 468
E common lanceifolium 444
E common kelleyi 353
2000m
1000m
0m
Mexico
Costa Rica
Panama
Colombia
Ecuador
Peru
French Guiana
Argentina
Enckea
Isophyllon
Macrostachys
Peltobryon
Pothomorphe
Radula
Schilleria
Unknown

Figure 4

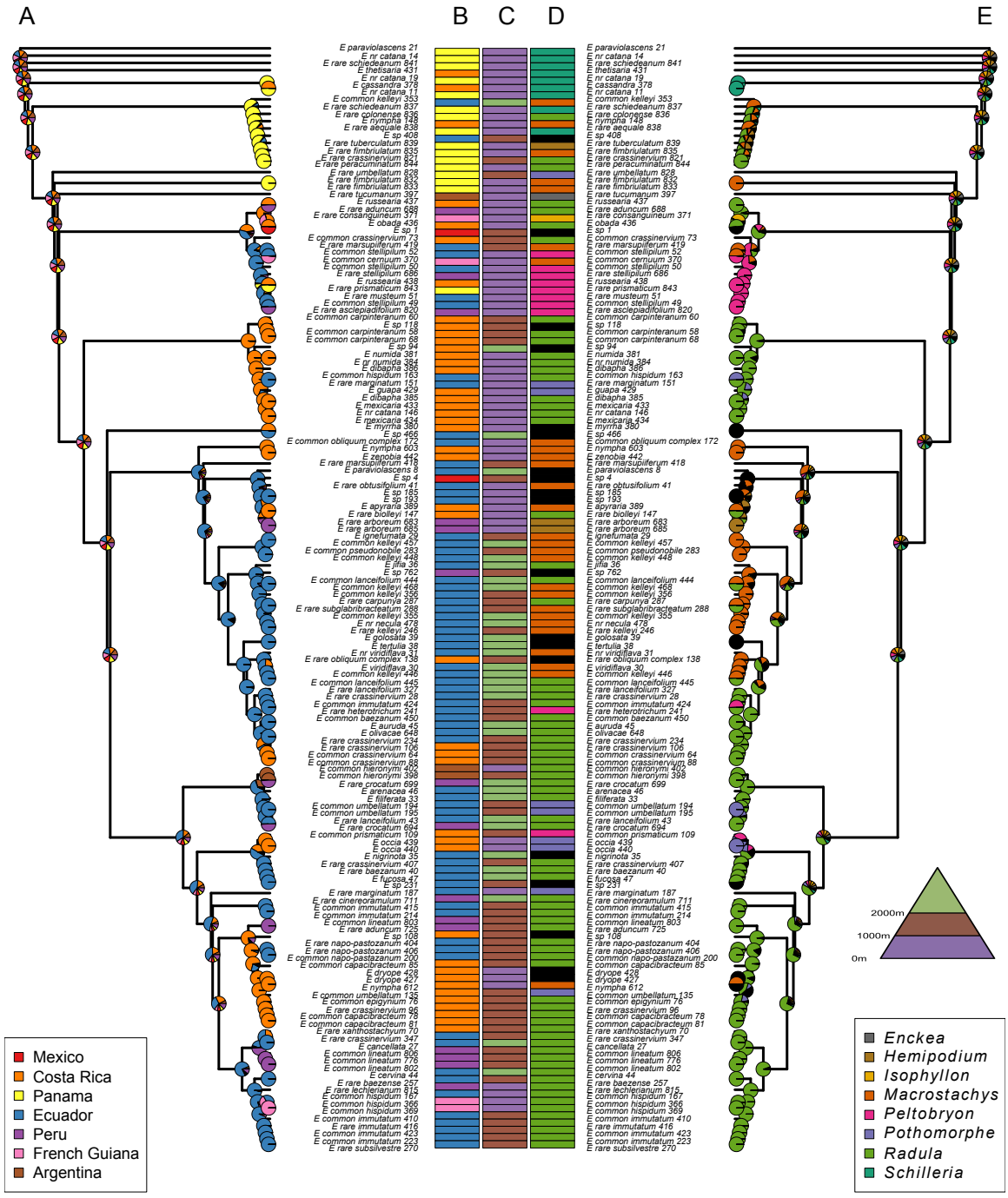

A
B
C
D
E
Mexico
Costa Rica
Panama
Ecuador
Peru
French Guiana
Argentina
2000m
1000m
0m
Enckea
Hemipodium
Isophyllon
Macrostachys
Peltobryon
Pothomorphe
Radula
Schilleria

Figure 5

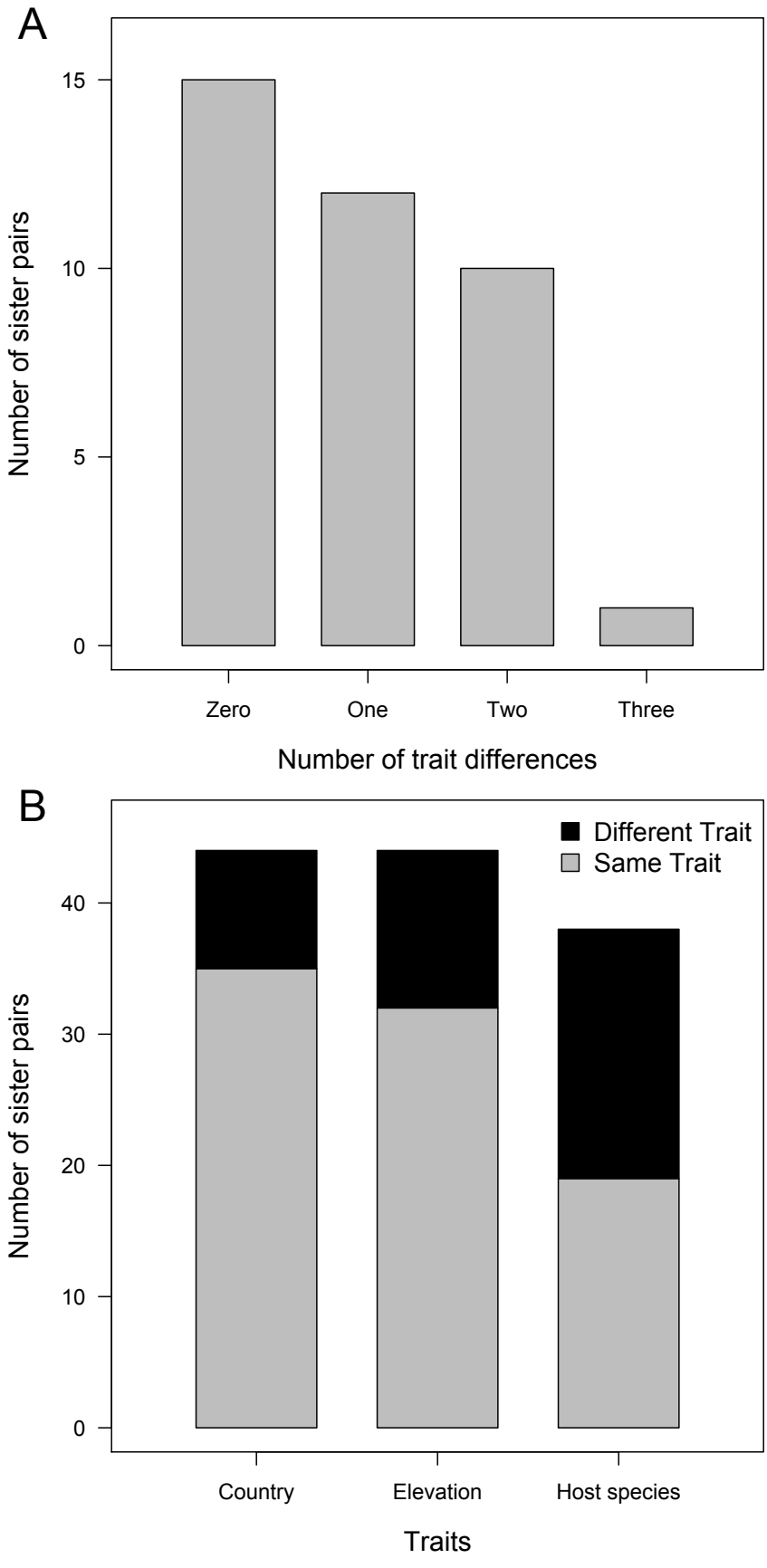
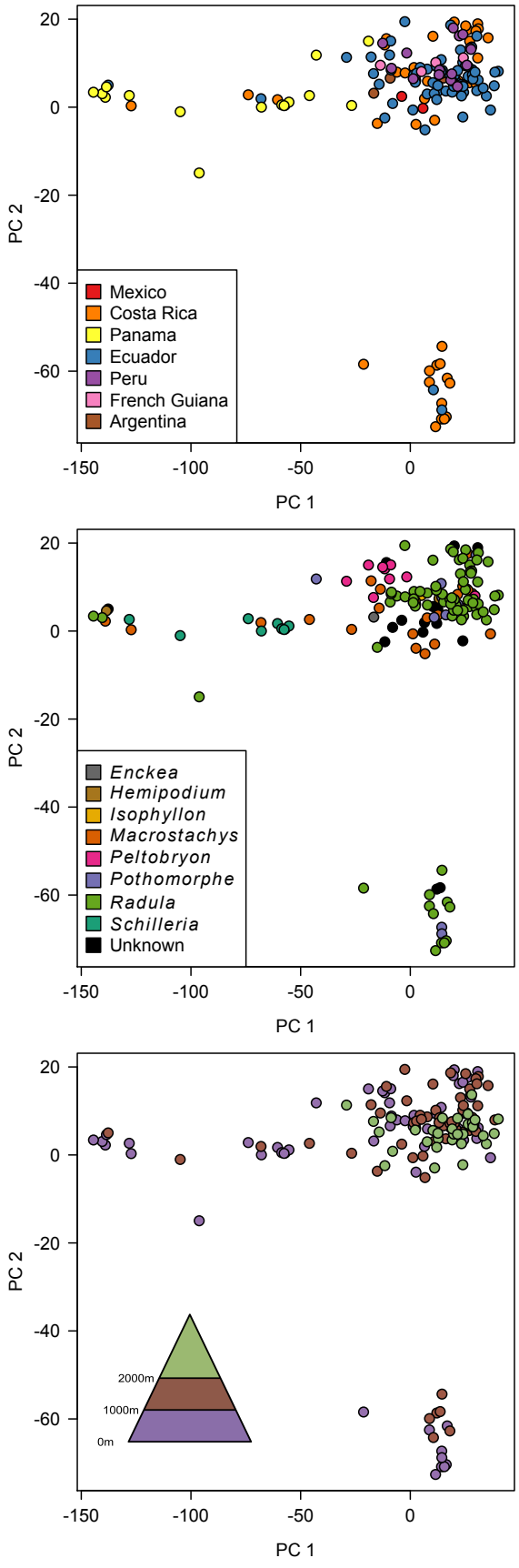

A
Number of sister pairs
0
5
10
15
Zero
One
Two
Three
Number of trait differences
B
Different Trait
Same Trait
Number of sister pairs
0
10
20
30
40
Country
Elevation
Host species
Traits

Figure 6


20
0
-20
-40
-60
PC 2
Mexico
Costa Rica
Panama
Ecuador
Peru
French Guiana
Argentina
-150
-100
-50
0
PC 1
20
0
-20
-40
-60
PC 2
Enckea
Hemipodium
Isophyllon
Macrostachys
Peltobryon
Pothomorphe
Radula
Schilleria
Unknown
-150
-100
-50
0
PC 1
20
0
-20
-40
-60
PC 2
2000m
1000m
0m
-150
-100
-50
0
PC 1

Supplementary Material

Chapter 1 ~ Host conservatism, geography, and elevation in the evolution of a Neotropical moth radiation

Joshua P. Jahner, Matthew L. Forister, Angela M. Smilanich, Thomas L. Parchman, Joseph S. Wilson, Eric J. Tepe, Lora A. Richards, Mario Alberto Quijano-Abril, Christopher S. Jeffrey, Andrea E. Glassmire, Lee A. Dyer

Table S1. Collection data for each *Eois* individual included in this study, including collection locality, collection country, elevation, host plant species, and host plant clade. Individual labels match the labels found in Figs. 3 & 4 in the main text. For each analysis (Sanger and GBS), included individuals are marked with a "1", while excluded individuals are marked with a "0".

Label	Species	Sanger	GBS	Country	location	Elevation (m)	Host plant	Host clade
E_sp_1	*Eois* sp.	0	1	Mexico	Varacruz, Instituto de ecologia AC	1300	Unknown	Unknown
E_sp_4	*Eois* sp.	1	1	Mexico	Varacruz, finca aluja zimpizahua	1100	Unknown	Unknown
E_paraviolascens_7	*Eois paraviolascens*	1	0	Ecuador	Yanayacu	2163	Unknown	Unknown
E_paraviolascens_8	*Eois paraviolascens*	1	1	Ecuador	Yanayacu	2163	Unknown	Unknown
E_nr_catana_11	*Eois* nr. *catana*	1	1	Panama	Barro Colorado Island	85	*Piper aequale*	*Schilleria*
E_nr_catana_14	*Eois* nr. *catana*	0	1	Panama	Barro Colorado Island	85	*Piper aequale*	*Schilleria*
E_nr_catana_19	*Eois* nr. *catana*	1	1	Panama	Barro Colorado Island	85	*Piper schiedeanum*	*Schilleria*
E_paraviolascens_21	*Eois paraviolascens*	1	1	Panama	Barro Colorado Island	85	*Piper aequale*	*Schilleria*
E_paraviolascens_22	*Eois paraviolascens*	1	0	Panama	Barro Colorado Island	85	*Piper aequale*	*Schilleria*
E_cancellata_26	*Eois cancellata*	1	0	Ecuador	Yanayacu	2163	*Piper schupii*	*Radula*
E_cancellata_27	*Eois cancellata*	1	1	Ecuador	Yanayacu	2163	*Piper schupii*	*Radula*
E_rare_crassinervium_28	*Eois* sp.	0	1	Ecuador	Yanayacu	2163	*Piper crassinervium*	*Radula*
E_ignefumata_29	*Eois ignefumata*	0	1	Ecuador	Las Palmas	1851	*Piper kelleyi*	*Macrostachys*
E_viridiflava_30	*Eois viridiflava*	0	1	Ecuador	Mirador	2428	*Piper kelleyi*	*Macrostachys*
E_nr_viridiflava_31	*Eois* nr. *viridiflava*	1	1	Ecuador	Sierra Azul	2231	*Piper kelleyi*	*Macrostachys*
E_filiferata_33	*Eois filiferata*	1	1	Ecuador	Yanayacu	2163	*Piper lanceifolium*	*Radula*
E_nigrinota_35	*Eois nigrinota*	0	1	Ecuador	Yanayacu	2163	*Piper* cf. *brevispicum*	Unknown
E_jifia_36	*Eois jifia*	0	1	Ecuador	Yanayacu	2163	*Piper perareolatum*	*Radula*
E_tertulia_38	*Eois tertulia*	0	1	Ecuador	Yanayacu	2163	Unknown	Unknown
E_golosata_39	*Eois golosata*	0	1	Ecuador	Yanayacu	2163	Unknown	Unknown
E_rare_baezanum_40	*Eois* sp.	1	1	Ecuador	Yanayacu	2163	*Piper baezanum*	*Radula*
E_rare_obtusifolium_41	*Eois* sp.	1	1	Ecuador	Jatun Sacha	414	*Piper obtusifolium*	*Macrostachys*
E_rare_lanceifolium_43	*Eois* sp.	1	1	Ecuador	Yanayacu	2163	*Piper lanceifolium*	*Radula*
E_cervina_44	*Eois cervina*	1	1	Ecuador	Yanayacu	2163	*Piper perareolatum*	*Radula*
E_auruda_45	*Eois auruda*	1	1	Ecuador	Yanayacu	2163	*Piper baezanum*	*Radula*
E_arenacea_46	*Eois arenacea*	1	1	Ecuador	Yanayacu	2163	*Piper lacunosum*	*Radula*
E_fucosa_47	*Eois fucosa*	1	1	Ecuador	Yanayacu	2163	*Piper perareolatum*	*Radula*
E_common_stellipilum_49	*Eois* sp.	1	1	Ecuador	Shiripuno	200	*Piper stellipilum*	*Peltobryon*
E_common_stellipilum_50	*Eois* sp.	1	1	Ecuador	Shiripuno	200	*Piper stellipilum*	*Peltobryon*
E_rare_musteum_51	*Eois* sp.	1	1	Ecuador	Shiripuno	200	*Piper musteum*	*Peltobryon*
E_common_stellipilum_52	*Eois* sp.	1	1	Ecuador	Shiripuno	200	*Piper stellipilum*	*Peltobryon*
E_common_carpinteranum_58	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Arboles Caidos	1291	*Piper carpinteranum*	*Radula*
E_common_carpinteranum_60	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Arboles Caidos	1291	*Piper carpinteranum*	*Radula*
E_common_crassinervium_64	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Arboles Caidos	1291	*Piper crassinervium*	*Radula*
E_common_carpinteranum_68	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Arboles Caidos	1291	*Piper carpinteranum*	*Radula*
E_rare_xanthostachyum_70	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Arboles Caidos	1291	*Piper xanthostachyum*	*Radula*
E_common_crassinervium_73	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper crassinervium*	*Radula*
E_sp_75	*Eois* sp.	1	0	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper* sp.	Unknown
E_common_epigynium_76	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper epigynium*	*Radula*
E_common_capacibracteum_78	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper capacibracteum*	*Radula*
E_common_epigynium_79	*Eois* sp.	1	0	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper epigynium*	*Radula*
E_common_capacibracteum_81	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper capacibracteum*	*Radula*
E_common_capacibracteum_85	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper capacibracteum*	*Radula*

E_common_crassinervium_88	*Eois* sp.	0	1	Costa Rica	Tapanti - Quinta St. Cecilia	1765	*Piper crassinervium*	*Radula*
E_sp_94	*Eois* sp.	1	1	Costa Rica	Tapanti - Esperanza	2627	*Piper* sp.	Unknown
E_sp_95	*Eois* sp.	1	0	Costa Rica	Tapanti - Esperanza	2627	*Piper* sp.	Unknown
E_rare_crassinervium_96	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper crassinervium*	*Radula*
E_rare_crassinervium_106	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper crassinervium*	*Radula*
E_sp_108	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper* sp.	Unknown
E_common_prismaticum_109	*Eois* sp.	1	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper prismaticum*	*Peltobryon*
E_sp_118	*Eois* sp.	0	1	Costa Rica	Tapanti - Sendero Oropendola	1291	*Piper* sp.	Unknown
E_common_umbellatum_123	*Eois* sp.	1	0	Costa Rica	Tapanti	1287	*Piper umbullatum*	*Pothomorphe*
E_common_umbellatum_124	*Eois* sp.	1	0	Costa Rica	Tapanti	1287	*Piper umbullatum*	*Pothomorphe*
E_common_umbellatum_135	*Eois* sp.	1	1	Costa Rica	Tapanti	1287	*Piper umbullatum*	*Pothomorphe*
E_rare_obliquum_complex_138	*Eois* sp.	0	1	Costa Rica	Tapanti - Coffee Plantation	1300	*Piper obliquum* complex	*Macrostachys*
E_nr_catana_146	*Eois* sp.	0	1	Costa Rica	Huertos	50	*Piper umbricola*	*Radula*
E_rare_biolleyi_147	*Eois* sp.	0	1	Costa Rica	La Selva	50	*Piper biolleyi*	*Radula*
E_nympha_148	*Eois nympha*	0	1	Costa Rica	La Selva	50	*Piper cenocladum*	*Macrostachys*
E_rare_marginatum_151	*Eois* sp.	0	1	Ecuador	Rio Palenque Reserve Sendero 1	200	*Piper marginatum*	*Pothomorphe*
E_common_hispidum_163	*Eois* sp.	0	1	Ecuador	Rio Palenque Reserve Sendero 1	190	*Piper hispidum*	*Radula*
E_common_hispidum_167	*Eois* sp.	1	1	Ecuador	Rio Palenque Reserve Sendero 1	190	*Piper hispidum*	*Radula*
E_common_obliquum_complex_172	*Eois* sp.	0	1	Ecuador	Rio Palenque Reserve Sendero 1	171	*Piper obliquum* complex	*Macrostachys*
E_sp_185	*Eois* sp.	1	1	Ecuador	Rio Palenque Reserve Sendero 3	187	*Piper* sp.	Unknown
E_rare_marginatum_187	*Eois* sp.	1	1	Ecuador	Rio Palenque Reserve Sendero 3	203	*Piper marginatum*	*Pothomorphe*
E_sp_193	*Eois* sp.	1	1	Ecuador	Rio Palenque Reserve Sendero 3	203	*Piper* sp.	Unknown
E_common_umbellatum_194	*Eois* sp.	0	1	Ecuador	Rio Anzu Reserve, Mera	1296	*Piper umbellatum*	*Pothomorphe*
E_common_umbellatum_195	*Eois* sp.	0	1	Ecuador	Rio Anzu Reserve, Mera	1296	*Piper umbellatum*	*Pothomorphe*
E_common_napo-pastazanum_200	*Eois* sp.	1	1	Ecuador	Rio Anzu Reserve, Mera	1296	*Piper napo-pastazanum*	*Radula*
E_common_immutatum_214	*Eois* sp.	0	1	Ecuador	Rio Anzu Reserve, Mera	1275	*Piper immutatum*	*Radula*
E_common_immutatum_223	*Eois* sp.	1	1	Ecuador	Zuñag Cientifo Estacion	1207	*Piper immutatum*	*Radula*
E_sp_231	*Eois* sp.	0	1	Ecuador	Zuñag Cientifo Estacion	1206	*Piper* sp.	Unknown
E_rare_crassinervium_234	*Eois* sp.	1	1	Ecuador	Zuñag Cientifo Estacion	1329	*Piper crassinervium*	*Radula*
E_rare_heterotrichum_241	*Eois* sp.	0	1	Ecuador	Zuñag Cientifo Estacion	1346	*Piper heterotrichum*	*Peltobryon*
E_sp_244	*Eois* sp.	1	0	Ecuador	Zuñag Cientifo Estacion	1346	*Piper* sp.	Unknown
E_rare_kelleyi_246	*Eois* sp.	0	1	Ecuador	Zuñag Cientifo Estacion	1581	*Piper kelleyi*	*Macrostachys*
E_rare_baezense_257	*Eois* sp.	1	1	Ecuador	Zuñag Cientifo Estacion	1581	*Piper baezense*	*Radula*
E_rare_subsilvestre_270	*Eois* sp.	0	1	Ecuador	Zuñag Cientifo Estacion	1564	*Piper subsilvestre*	*Radula*
E_common_pseudonobile_283	*Eois* sp.	1	1	Ecuador	Mindo	1332	*Piper pseudonobile*	*Macrostachys*
E_rare_carpunya_287	*Eois* sp.	1	1	Ecuador	Mindo	1366	*Piper carpunya*	*Radula*
E_rare_subglabribracteatum_288	*Eois* sp.	1	1	Ecuador	Mindo	1366	*Piper subglabribracteatum*	*Macrostachys*
E_rare_lanceifolium_327	*Eois* sp.	1	1	Ecuador	Los Angeles en Mira Flores, Mindo	2002	*Piper lanceifolium*	*Radula*
E_rare_crassinervium_347	*Eois* sp.	1	1	Ecuador	Los Angeles en Mira Flores, Mindo	1984	*Piper crassinervium*	*Radula*
E_common_pseudonobile_349	*Eois* sp.	1	0	Ecuador	Los Angeles en Mira Flores, Mindo	1842	*Piper pseudonobile*	*Macrostachys*
E_common_kelleyi_350	*Eois* sp.	1	0	Ecuador	Las Palmas	2086	*Piper kelleyi*	*Macrostachys*
E_common_kelleyi_353	*Eois* sp.	1	1	Ecuador	Guacamayos	2097	*Piper kelleyi*	*Macrostachys*
E_common_kelleyi_355	*Eois* sp.	1	1	Ecuador	Guacamayos	2097	*Piper kelleyi*	*Macrostachys*
E_common_kelleyi_356	*Eois* sp.	1	1	Ecuador	Cocodrillos	1811	*Piper kelleyi*	*Macrostachys*
E_common_kelleyi_360	*Eois* sp.	1	0	Ecuador	Mirador	2255	*Piper kelleyi*	*Macrostachys*
E_common_hispidum_366	*Eois* sp.	1	1	French Guiana	Estación B. Nouragues	100	*Piper hispidum*	*Radula*

E_common_hispidum_369	*Eois* sp.	1	1	French Guiana	Estación B. Nouragues	100	*Piper hispidum*	*Radula*
E_common_cernuum_370	*Eois* sp.	0	1	French Guiana	Estación B. Nouragues	100	*Piper cernuum*	*Macrostachys*
E_rare_consanguineum_371	*Eois* sp.	1	1	French Guiana	Estación B. Nouragues	100	*Piper consanguineum*	*Isophyllon*
E_common_anonifolium_374	*Eois* sp.	1	0	French Guiana	Estación B. Nouragues	100	*Piper anonifolium*	*Isophyllon*
E_common_cernuum_375	*Eois* sp.	1	0	French Guiana	Estación B. Nouragues	100	*Piper cernuum*	*Macrostachys*
E_cassandra_378	*Eois cassandra*	1	1	Costa Rica	Huertos	50	*Piper schiedeanum*	*Schilleria*
E_myrrha_380	*Eois myrrha*	0	1	Costa Rica	La Selva	50	Unknown	Unknown
E_numida_381	*Eois numida*	0	1	Costa Rica	Huertos	50	*Piper hispidum*	*Radula*
E_nr_numida_384	*Eois* nr. *numida*	1	1	Costa Rica	La Selva	50	*Piper hispidum*	*Radula*
E_dibapha_385	*Eois dibapha*	0	1	Costa Rica	La Selva	50	*Piper umbricola*	*Radula*
E_dibapha_386	*Eois dibapha*	1	1	Costa Rica	La Selva	50	*Piper umbricola*	*Radula*
E_apyraria_389	*Eois apyraria*	0	1	Costa Rica	La Selva	50	*Piper evasum*	*Macrostachys*
E_rare_tucumanum_397	*Eois* sp.	1	1	Argentina	Tucuman Site 1	450	*Piper tucumanum*	*Enckea*
E_common_hieronymi_398	*Eois* sp.	1	1	Argentina	Tucuman Site 2	1400	*Piper hieronymi*	*Radula*
E_common_hieronymi_402	*Eois* sp.	1	1	Argentina	Tucuman Site 4	800	*Piper hieronymi*	*Radula*
E_rare_napo-pastozanum_404	*Eois* sp.	0	1	Ecuador	Narupa - Sendero Chonta Yacu	1186	*Piper napo-pastazanum*	*Radula*
E_rare_napo-pastozanum_406	*Eois* sp.	0	1	Ecuador	Narupa - Sendero Chonta Yacu	1186	*Piper napo-pastazanum*	*Radula*
E_rare_crassinervium_407	*Eois* sp.	1	1	Ecuador	Narupa - Sendero Chonta Yacu	1186	*Piper crassinervium*	*Radula*
E_sp_408	*Eois* sp.	0	1	Ecuador	Narupa	1186	*Piper* sp.	Unknown
E_common_immutatum_410	*Eois* sp.	0	1	Ecuador	Narupa - Sendero de Jibaro	1186	*Piper immutatum*	*Radula*
E_common_immutatum_415	*Eois* sp.	1	1	Ecuador	Narupa	1186	*Piper immutatum*	*Radula*
E_rare_immutatum_416	*Eois* sp.	0	1	Ecuador	Narupa	1186	*Piper immutatum*	*Radula*
E_rare_marsupiiferum_418	*Eois* sp.	0	1	Ecuador	Narupa	1186	*Piper marsupiiferum*	*Macrostachys*
E_rare_marsupiiferum_419	*Eois* sp.	1	1	Ecuador	Narupa	1186	*Piper marsupiiferum*	*Macrostachys*
E_common_immutatum_423	*Eois* sp.	0	1	Ecuador	Rio Narupa	1186	*Piper immutatum*	*Radula*
E_common_immutatum_424	*Eois* sp.	0	1	Ecuador	Rio Narupa	1186	*Piper immutatum*	*Radula*
E_dryope_427	*Eois dryope*	1	1	Costa Rica	Huertos	50	*Piper* sp.	Unknown
E_dryope_428	*Eois dryope*	1	1	Costa Rica	Huertos	50	*Piper* sp.	Unknown
E_guapa_429	*Eois guapa*	0	1	Costa Rica	La Selva	50	*Piper peltatum*	*Pothomorphe*
E_thetisaria_431	*Eois thetisaria*	0	1	Costa Rica	Huertos	50	*Piper schiedeanum*	*Schilleria*
E_mexicaria_433	*Eois mexicaria*	1	1	Costa Rica	La Selva	50	*Piper umbricola*	*Radula*
E_mexicaria_434	*Eois mexicaria*	1	1	Costa Rica	La Selva	50	*Piper umbricola*	*Radula*
E_obada_436	*Eois obada*	0	1	Costa Rica	La Selva	50	*Piper sublineatum*	*Radula*
E_russearia_437	*Eois russearia*	1	1	Costa Rica	La Selva	50	*Piper hispidum*	*Radula*
E_russearia_438	*Eois russearia*	0	1	Costa Rica	La Selva	50	*Piper machadoanum*	*Peltobryon*
E_occia_439	*Eois occia*	0	1	Costa Rica	Huertos	50	*Piper auritum*	*Pothomorphe*
E_occia_440	*Eois occia*	0	1	Costa Rica	Huertos	50	*Piper auritum*	*Pothomorphe*
E_zenobia_442	*Eois zenobia*	0	1	Costa Rica	Tirimbina	80	*Piper cenocladum*	*Macrostachys*
E_common_lanceifolium_444	*Eois* sp.	1	1	Ecuador	Sierra Azul	2231	*Piper lanceifolium*	*Radula*
E_common_lanceifolium_445	*Eois* sp.	1	1	Ecuador	Sierra Azul	2231	*Piper lanceifolium*	*Radula*
E_common_kelleyi_446	*Eois* sp.	0	1	Ecuador	Yanayacu	2163	*Piper kelleyi*	*Macrostachys*
E_common_kelleyi_448	*Eois* sp.	1	1	Ecuador	Guacamayos	2097	*Piper kelleyi*	*Macrostachys*
E_common_baezanum_450	*Eois* sp.	1	1	Ecuador	Cocodrillos	1811	*Piper baezanum*	*Radula*
E_common_kelleyi_457	*Eois* sp.	1	1	Ecuador	Yanayacu	2163	*Piper kelleyi*	*Macrostachys*
E_sp_466	*Eois* sp.	1	1	Ecuador	Yanayacu	2163	Unknown	Unknown
E_common_kelleyi_468	*Eois* sp.	1	1	Ecuador	Yanayacu	2163	*Piper kelleyi*	*Macrostachys*

E_nr_necula_478	*Eois* nr. *necula*	1	1	Ecuador	Yanayacu	2163	*Piper kelleyi*	*Macrostachys*
E_nympha_603	*Eois nympha*	1	1	Costa Rica	La Selva	50	*Piper cenocladum*	*Macrostachys*
E_nympha_612	*Eois nympha*	1	1	Costa Rica	La Selva	50	*Piper cenocladum*	*Macrostachys*
E_olivacae_648	*Eois olivacea*	1	1	Ecuador	Yanayacu	2163	*Piper crassinervium*	*Radula*
E_rare_arboreum_683	*Eois* sp.	0	1	Peru	Jenaro Herrera Research Center	91	*Piper arboreum*	*Hemipodium*
E_rare_arboreum_685	*Eois* sp.	0	1	Peru	Jenaro Herrera Research Center	91	*Piper arboreum*	*Hemipodium*
E_rare_stellipilum_686	*Eois* sp.	0	1	Peru	Jenaro Herrera Research Center	121	*Piper stellipilum*	*Peltobryon*
E_rare_aduncum_688	*Eois* sp.	1	1	Peru	Jenaro Herrera Research Center	114	*Piper aduncum*	*Radula*
E_rare_crocatum_694	*Eois* sp.	0	1	Peru	Wayquecha BS	2981	*Piper crocatum*	*Radula*
E_rare_crocatum_699	*Eois* sp.	1	1	Peru	Wayquecha BS	2960	*Piper crocatum*	*Radula*
E_rare_cinereoramulum_711	*Eois* sp.	0	1	Peru	Manu Road	2220	*Piper cinereoramulum*	*Radula*
E_rare_aduncum_725	*Eois* sp.	0	1	Peru	Posada San Pedro Lodge	1398	*Piper aduncum*	*Radula*
E_sp_762	*Eois* sp.	1	1	Peru	Posada San Pedro Lodge	1382	*Piper* sp.	Unknown
E_common_lineatum_776	*Eois* sp.	1	1	Peru	Posada San Pedro Lodge	1428	*Piper lineatum*	*Radula*
E_common_lineatum_802	*Eois* sp.	0	1	Peru	Posada San Pedro Lodge	1428	*Piper lineatum*	*Radula*
E_common_lineatum_803	*Eois* sp.	0	1	Peru	Posada San Pedro Lodge	1428	*Piper lineatum*	*Radula*
E_common_lineatum_806	*Eois* sp.	1	1	Peru	Posada San Pedro Lodge	1428	*Piper lineatum*	*Radula*
E_rare_lechlerianum_815	*Eois* sp.	1	1	Peru	Salvacion	491	*Piper lechlerianum*	*Radula*
E_rare_asclepiadifolium_820	*Eois* sp.	0	1	Peru	Salvacion	482	*Piper asclepiadifolium*	*Peltobryon*
E_rare_crassinervium_821	*Eois* sp.	0	1	Panama	Estacion Fortuna	1265	*Piper crassinervium*	*Radula*
E_rare_umbellatum_828	*Eois* sp.	1	1	Panama	Parque Nac Baru	1900	*Piper umbullatum*	*Pothomorphe*
E_rare_fimbriulatum_832	*Eois* sp.	1	1	Panama	Cerro Campana	820	*Piper fimbriulatum*	*Macrostachys*
E_rare_fimbriulatum_833	*Eois* sp.	1	1	Panama	Cerro Campana	820	*Piper fimbriulatum*	*Macrostachys*
E_rare_fimbriulatum_835	*Eois* sp.	0	1	Panama	Cerro Campana	820	*Piper fimbriulatum*	*Macrostachys*
E_rare_colonense_836	*Eois* sp.	0	1	Panama	Parque Soberania	130	*Piper colonense*	*Radula*
E_rare_schiedeanum_837	*Eois* sp.	0	1	Panama	Parque Soberania	130	*Piper schiedeanum*	*Schilleria*
E_rare_aequale_838	*Eois* sp.	0	1	Panama	Parque Soberania	130	*Piper aequale*	*Schilleria*
E_rare_tuberculatum_839	*Eois* sp.	0	1	Panama	Sendero Laguna, Gamboa	45	*Piper tuberculatum*	*Hemipodium*
E_rare_schiedeanum_841	*Eois* sp.	0	1	Panama	Parque Soberania	130	*Piper schiedeanum*	*Schilleria*
E_rare_prismaticum_843	*Eois* sp.	1	1	Panama	Llano Carti	325	*Piper prismaticum*	*Peltobryon*
E_rare_peracuminatum_844	*Eois* sp.	0	1	Panama	Santa Rita	250	*Piper peracuminatum*	*Radula*
E_rare_salgaranum_846	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper salgaranum*	*Radula*
E_rare_auritum_847	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper auritum*	*Pothomorphe*
E_rare_scobinifolium_866	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper scobinifolium*	*Radula*
E_rare_crassinervium_869	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper crassinervium*	*Radula*
E_rare_scobinifolium_880	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper scobinifolium*	*Radula*
E_rare_haughtii_891	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper haughtii*	*Radula*
E_rare_gorgonillense_898	*Eois* sp.	1	0	Colombia	Cocorna	1300	*Piper gorgonillense*	*Peltobryon*
E_rare_gorgonillense_951	*Eois* sp.	1	0	Colombia	San Luis	700	*Piper gorgonillense*	*Peltobryon*
E_rare_haughtii_956	*Eois* sp.	1	0	Colombia	San Luis	700	*Piper haughtii*	*Radula*
E_rare_aduncum_993	*Eois* sp.	1	0	Colombia	San Luis	700	*Piper aduncum*	*Radula*
E_rare_aduncum_1002	*Eois* sp.	1	0	Colombia	San Luis	700	*Piper aduncum*	*Radula*

Table S2. PCR primers for each gene are listed. Wingless and EF1-alpha primers have universal sequencing primers attached on the ends to facilitate sequencing.

Gene	Primer	Sequence	Reference
COI	EoJoe	ATT AAT TCG AGC WGA AYT AGG	This study
	LepR	TAA ACT TCT GGA TGT CCA AA	Sheffield et al. 2009
EF1- α	HybAlF	TAA TAC GAC TCA CTA TAG GGG AGG AAA TYA ARA ArG AAG	Wahlberg & Wheat 2008
	HybEFrcM4	ATT AAC CCT CAC TAA AGA CAG CVA CKG TYT GYC TCA TRT C	Wahlberg & Wheat 2008
Wingless	HybLepWG1	TAA TAC GAC TCA CTA TAG GGG ART GYA ART GYC AYG GYA TGT CTG G	Wahlberg & Wheat 2008
	HybLepWG2	ATT AAC CCT CAC TAA AGA CTI CGC ARC ACC ART GGA ATG TRC A	Wahlberg & Wheat 2008
Sequencing Primers	T7Promoter(F)	TAA TAC GAC TCA CTA TAG GG	Wahlberg & Wheat 2008
	T3(R)	ATT AAC CCT CAC TAA AG	Wahlberg & Wheat 2008

Table S3. Summary of PCR protocols and master mix recipe for the three genes sequenced in this study.

	COI	EF1- α	WG
DNA Volume	1μL	4μL	4μL
H_20 Volume	17.5μL	7.64μL	9.5μL
dNTP Volume	1μL	0.6μL	0.4μL
10X buffer Volume	2.5μL	1.5μL	2μL
$MgCl_2$ Volume	0μL	0μL	2μL
Primer 1 Volume	1μL	0.5μL	1μL
Primer 2 Volume	1μL	0.5μL	1μL
Taq Volume	1μL	0.26μL	0.1μL
Initialization time	3 min	2 min	3 min
Initialization temp	94°C	94°C	94°C
Denaturation time	30s	30s	30s
Denaturation temp	94°C	94°C	94°C
Annealing time	60s	30s	30s
Annealing temp	50°C	55°C	62°C
Elongation time	60s	90s	90s
Elongation temp	72°C	72°C	72°C
Final Elongation time	10 min	30 min	10 min
Final Elongation temp	72°C	72°C	72°C
Number of Cycles	40	40	40

Fig. S1

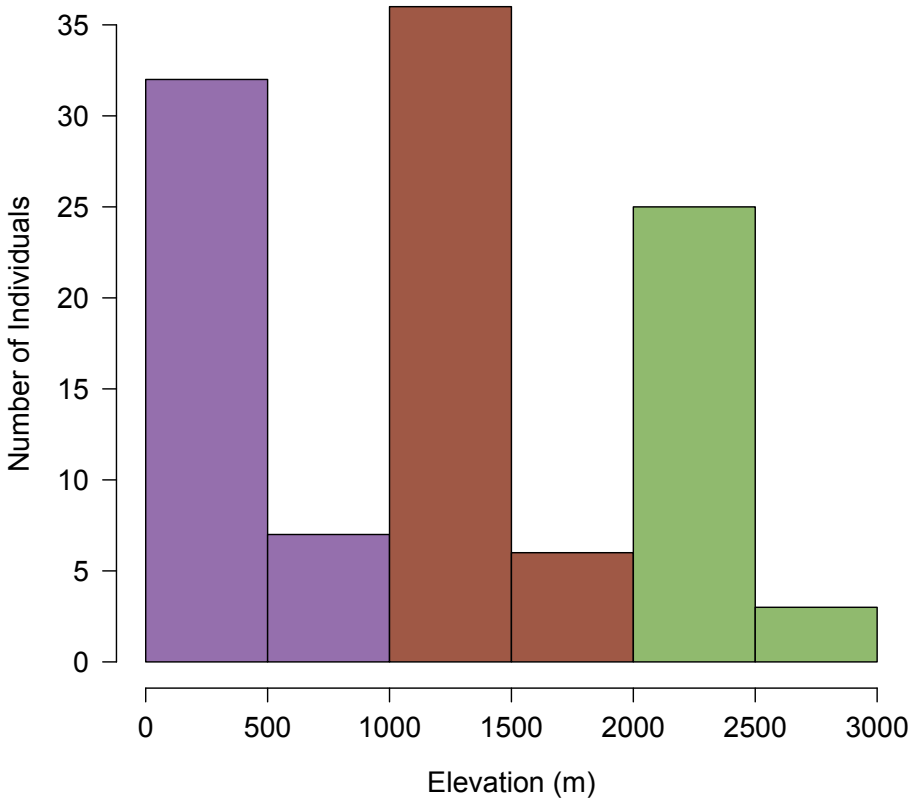


Fig. S1. A histogram displaying the distribution of *Eois* individuals collected across elevation. For categorical trait analyses, individuals were assigned to one of three categories: <1,000m (purple); 1,000-2,000m (brown); >2,000m (green).

Fig. S2

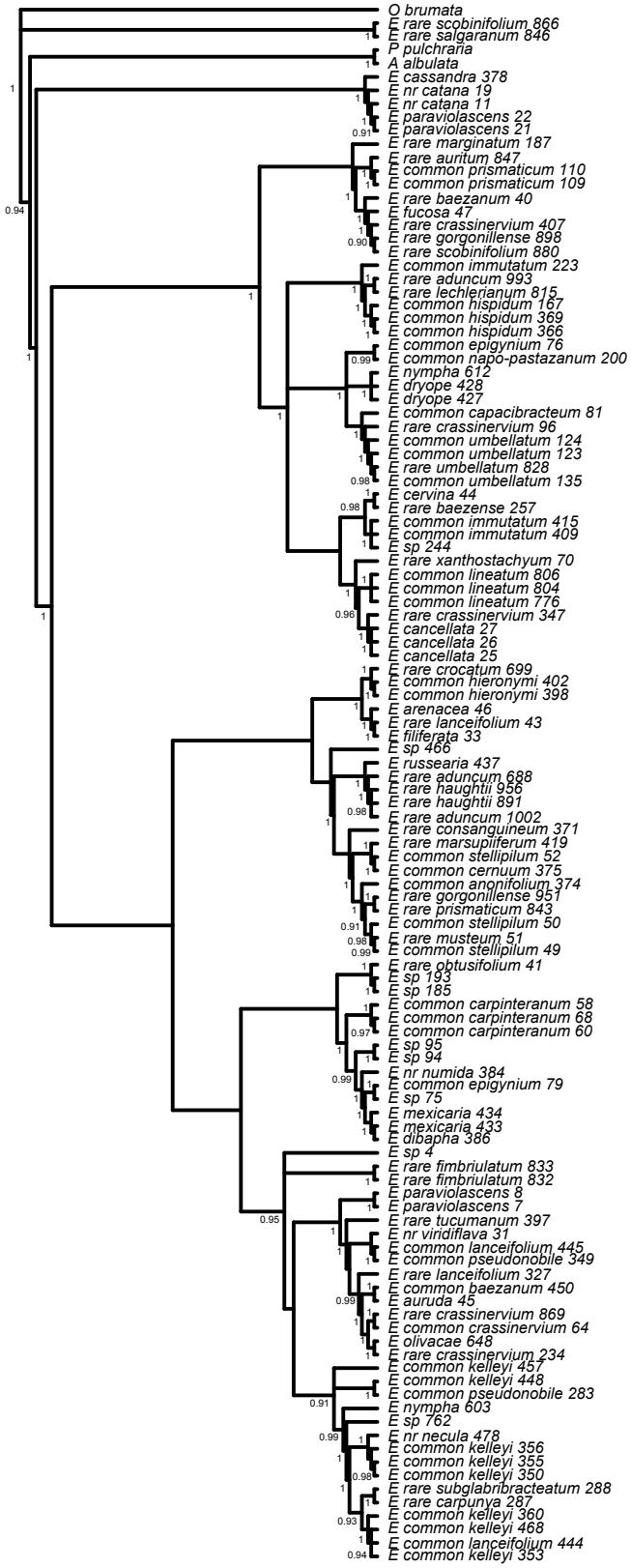

O brumata
E rare scobinifolium 866
E rare salgaranum 846
P pulchraria
A albulata
E cassandra 378
E nr catana 19
E nr catana 11
E paraviolascens 22
E paraviolascens 21
E rare marginatum 187
E rare auritum 847
E common prismaticum 110
E common prismaticum 109
E rare baezanum 40
E fucosa 47
E rare crassinervium 407
E rare gorgonillense 898
E rare scobinifolium 880
E common immutatum 223
E rare aduncum 993
E rare lechlerianum 815
E common hispidum 167
E common hispidum 369
E common hispidum 366
E common epigynium 76
E common napo-pastazanum 200
E nympha 612
E dryope 428
E dryope 427
E common capacibracteum 81
E rare crassinervium 96
E common umbellatum 124
E common umbellatum 123
E rare umbellatum 828
E common umbellatum 135
E cervina 44
E rare baezense 257
E common immutatum 415
E common immutatum 409
E sp 244
E rare xanthostachyum 70
E common lineatum 806
E common lineatum 804
E common lineatum 776
E rare crassinervium 347
E cancellata 27
E cancellata 26
E cancellata 25
E rare crocatum 699
E common hieronymi 402
E common hieronymi 398
E arenacea 46
E rare lanceifolium 43
E filiferata 33
E sp 466
E russearia 437
E rare aduncum 688
E rare haughtii 956
E rare haughtii 891
E rare aduncum 1002
E rare consanguineum 371
E rare marsupiferum 419
E common stellipilum 52
E common cernuum 375
E common anomifolium 374
E rare gorgonillense 951
E rare prismaticum 843
E common stellipilum 50
E rare musteum 51
E common stellipilum 49
E rare obtusifolium 41
E sp 193
E sp 185
E common carpinteranum 58
E common carpinteranum 68
E common carpinteranum 60
E sp 95
E sp 94
E nr numida 384
E common epigynium 79
E sp 75
E mexicaria 434
E mexicaria 433
E dibapha 386
E sp 4
E rare fimbriulatum 833
E rare fimbriulatum 832
E paraviolascens 8
E paraviolascens 7
E rare tucumanum 397
E nr viridiflava 31
E common lanceifolium 445
E common pseudonobile 349
E rare lanceifolium 327
E common baezanum 450
E auruda 45
E rare crassinervium 869
E common crassinervium 64
E olivacae 648
E rare crassinervium 234
E common kelleyi 457
E common kelleyi 448
E common pseudonobile 283
E nympha 603
E sp 762
E nr necula 478
E common kelleyi 356
E common kelleyi 355
E common kelleyi 350
E rare subglabribracteatum 288
E rare carpunya 287
E common kelleyi 360
E common kelleyi 468
E common lanceifolium 444
E common kelleyi 353

Fig S2. The topology of the concatenated phylogeny (including the COI, EF1-α, and WG genes) is displayed. Individuals not identified to species are labeled as either rare or common (depending on how many individuals were collected), followed by the host plant species name. Node support values are listed for nodes with posterior probabilities >0.90. See Fig. 3 in the main text for ancestral trait reconstructions and the tip states for country, elevation, and host clade.

Supplementary References

Sheffield CS, Hebert PDN, Kevan PF, Packer L. 2009. DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. *Molecular Ecology Resources* 9:196-207.

Wahlberg N, Wheat CWW. 2008. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Systematic Biology* 57:231-242.

Chapter 2 ~ Phylogenetic signal and the evolution of secondary chemistry in *Radula* (Piperaceae): persistent trait correlations despite rampant trait lability

Joshua P. Jahner,[1,2] Eric J. Tepe,[3] Lora A. Richards,[1,2] Matthew L. Forister,[1,2] Lee A. Dyer,[1,2,4] Angela M. Smilanich,[1,2] Craig D. Dodson,[5] Christopher S. Jeffrey[1,5]

[1]Program in Ecology, Evolution, and Conservation Biology, University of Nevada, Reno, NV 89557, USA; [2]Department of Biology, University of Nevada, Reno, NV 89557, USA; [3]Department of Biological Sciences, University of Cincinnati, Cincinnati, OH 45221, USA; [4]Sección Invertebrados, Museo Ecuatoriano de Ciencias Naturales, Quito, Ecuador; [5]Department of Chemistry, University of Nevada, Reno, NV 89557, USA

KEYWORDS: Blomberg's *K*, codiversification, metabolomics, network, nuclear magnetic resonance, Pagel's λ, *Piper*, phylogenetic comparative analyses, plant-insect interaction, trait evolution

ABSTRACT

Ecological interactions between plants and the insects that feed on them are structured by secondary chemistry, which in turn shapes long-term evolutionary patterns. Hypotheses regarding the evolutionary interplay between plants and insects posit that closely related plants should have similar chemical profiles, a pattern referred to as phylogenetic signal. Previous studies have documented mixed strengths of phylogenetic signal for a variety of compounds across different plant lineages, perhaps indicating that the distribution of phylogenetic signal is contextual; however, most of these studies have focused on one or two compounds, and almost none have quantified secondary chemistry using a metabolomic approach. In this study, we generated metabolomic data from NMR spectroscopy to investigate the distribution of phylogenetic signal in a diverse clade of plants (*Radula* clade of *Piper*; Piperaceae) that is host to a number of specialized herbivores. NMR spectra were generated for 48 *Radula* specimens and the relationships among chemical shifts were used to delineate 14 broad chemical modules consistent with chemical structures. In general, the evolution of secondary chemistry is highly labile in *Radula*, with few traits displaying phylogenetic signal. Despite this, a number of significant correlations among modules were recovered, suggesting that while individual traits might not be conserved, groups of traits consistently co-occur across species. These results demonstrate the importance of interactions among alternative defense compounds in shaping the evolution of secondary chemistry within a rapidly multiplying lineage of well-defended plants.

INTRODUCTION

Secondary chemistry plays a pivotal role in structuring ecological interactions and driving evolutionary dynamics between plants and the insects that feed on them (Fraenkel 1959; Ehrlich & Raven 1964; Berenbaum 1983; Agrawal 2011). For instance, host chemistry mediates resource utilization in insects at a variety of stages throughout the interaction: leaf chemistry can shape adult oviposition preferences (Thompson & Pellmyr 1991); the presence of specific chemical compounds can stimulate larval feeding (Bowers 1983, 1984); and sequestered metabolites can decrease immune function against natural enemies (Smilanich *et al.* 2009, Richards et al. 2012). In addition, certain chemical metabolites act as highly effective selective pressures against herbivores that are not physiologically capable of tolerating them, resulting in a world dominated by a preponderance of specialist herbivores feeding on well-defended plants (Forister *et al*. 2015). As such, plants capable of developing novel chemical defenses that deter herbivory are hypothesized to accrue higher fitness in response to enemy release (e.g., Berenbaum 1978), potentially resulting in the diversification of a chemically well-defended lineage of plants (the escape and radiate hypothesis; Ehrlich & Raven 1964; Thompson 1989; reviewed by Janz 2011).

One expectation from an escape and radiate scenario is that closely related plants should have similar chemical profiles, presumably because all plant metabolites are built from a few biosynthetic pathways (Agrawal 2007). This pattern, wherein closely related taxa share more similar trait values than more distant relatives, is referred to as phylogenetic signal (Blomberg & Garland, 2002). The advent of phylogenetic comparative methods specifically designed to quantify phylogenetic signal (e.g., Pagel

1999; Blomberg *et al*. 2003; reviewed by Münkemüller *et al.* 2012; Kamilar & Cooper 2013) has allowed researchers to test hypotheses regarding the evolution of secondary chemistry in a quantitative phylogenetic framework that is comparable across studies and taxa (Agrawal 2007). Overall, many studies have reported mixed strengths of phylogenetic signal in chemical traits, depending on the class of chemical compound (e.g., Agrawal *et al.* 2009; Johnson *et al*. 2009b, 2014; Kursar et al. 2009; Kariñho-Betancourt *et al.* 2015). For example, Kursar *et al.* (2009) found moderate phylogenetic signal in phenolics across the tropical tree genus *Inga* (Fabaceae), but not in saponins or amino/imino acids. In a similar vein, correlations in chemical traits among species have been documented in some lineages (e.g., Johnson *et al.* 2014; Kariñho-Betancourt *et al.* 2015), which are consistent with a hypothesis that phytochemical diversity is structured by suites of codiversifying traits or tradeoffs among traits (positive and negative correlations, respectively). Additionally, the bioprospecting field has recently made a strong push to use phylogenetic comparative methods when searching for groups of closely related plants that contain compounds of medicinal value, with a number of studies reporting phylogenetic signal in a variety of traits (Saslis-Lagoudakis *et al.* 2011, 2012; Rønsted *et al.* 2012; Yessoufou *et al.* 2015).

Recent breakthroughs in metabolomic approaches now allow for the characterization of entire suites of metabolites in plants, which is opening up a new frontier in our ability to address long-standing hypotheses on the evolution of phytochemistry, especially in the context of plant-herbivore interactions (Macel *et al.* 2010). Such approaches are vital when examining secondary defense chemistry because many secondary metabolites act synergistically, where mixtures of compounds have

larger effects on herbivores than single compounds (Dyer *et al*. 2003; Richards *et al*. 2010). In addition, a focus on single compounds is likely insufficient when plants have compounds with redundant but complementary functions (Jones & Firn 1991; Romeo *et al.* 1996). In this study, we utilized nuclear magnetic resonance (NMR) spectroscopy in a metabolomic framework (Krishnan *et al.* 2005) to examine the evolution of chemistry in *Piper* (Piperaceae), a hyper-diverse genus of tropical plants with a rich history of studies examining chemically mediated plant-insect interactions (e.g., Dyer *et al.* 2001, 2004b; Tepe *et al.* 2014; reviewed by Dyer & Palmer 2004). Specifically, the main objectives of this study were: 1) to generate and characterize metabolomic data for a large number of *Piper* species (specifically, from the *Radula* clade); 2) to quantify the distribution of phylogenetic signal across phytochemical traits; and 3) to determine if suites of chemical traits evolved in concert (positive correlations) or if chemical tradeoffs (negative correlations) are more prominent in structuring chemical profiles.

METHODS

Study system and sample collection

Piper is known for being highly chemically diverse (Richards *et al*. in review), especially in alkaloid compounds (Gutierrez *et al.* 2013) that are known to deter herbivory (Dyer et al. 2003). To date, 112 *Piper* species worldwide have been chemically profiled, yielding 667 different compounds distributed as follows: 190 alkaloids/amides, 49 lignans, 70 neolignans, 97 terpenes, 39 propenylphenols, 15 steroids, 18 kavapyrones, 17 chalcones/ dihydrochalcones, 16 flavones, 6 flavanones, 4 piperolides (cinnamylidone

butenolides) and 146 compounds that do not fit into the major categories of secondary metabolites (Parmar et al. 1997; Dyer et al. 2004a; Kato & Furlan 2007). Constitutive secondary metabolites have been found in all parts of the plant, and there are countless demonstrations of strong synergistic effects against herbivores and pathogens for all classes of *Piper* compounds, especially in the amides (Dyer et al. 2004a; Richards et al. 2010). Our main goal in this study was to sample a wide range of geographic, phylogenetic, and chemical variation within the clade *Radula* (see Table S1). We focused on the *Radula* clade due to the high species and chemical diversity found within the group. Overall we collected 48 *Radula* specimens from four countries: Brazil, Costa Rica, Ecuador, and Panama. For chemical and genetic profiling, we selected the youngest fully expanded leaves for collection and dried them immediately with silica gel. Voucher specimens were pressed, dried, and deposited in one or more herbaria for future reference and species verification (Table S1).

Chemical profiling

Leaf samples were ground to a fine powder and 2g were transferred to a screw cap test tube and combined with 10 ml of methanol. The samples were sonicated for 10 minutes and filtered to separate the leaf material from the supernatant. This step was repeated a second time and the supernatants were combined and transferred to a pre-weighed 20 ml scintillation vial. The solvent was removed under reduced pressure at 30°C and prepared for NMR analysis. The crude extracts were analyzed in dueterated methanol (CD3OD) and TMS on a Varian 400 MHz solution state NMR with autosampler. There are several advantages of using NMR in these types of comparisons

which include: the ability to observe a wide range of compound classes in a single analysis, non-destructive analysis, reproducibility, superior structural resolution, and simple peak alignment. The NMR data were processed using MestReNova software (Mestrelab Research, Spain). Spectra from the crude extracts were aligned using the solvent peak, baseline corrected, phase corrected, and binned every 0.04 ppm from 0.5 to 12 ppm. The solvent and water peaks were removed and the binned spectra were normalized to the total area of 100.

We analyzed binned spectral data using a weighted network approach (Horvath 2011) to identify clusters of chemical shifts that co-occur across the 48 samples for which we had genetic sequence data. Using the WGCNA package (Langfelder and Horvath 2008) in R v3.0.1 (R Core Team 2013), we first determined the appropriate power transformation (β) to use in calculating the adjacency of the network (Figure S2). In using a soft threshold, we calibrated β for the dataset to satisfy scale-free topology in the network. In terms of binned spectral data, this ensures that we can differentiate baseline values from meaningful peaks. From this analysis, we used soft thresholding power 9 to construct a network and identify clusters of co-occurring chemical shifts (modules), with a minimum module size of 3 chemical shifts, using the blockwiseModules function in WGCNA package in R. The network was visualized in Cytoscape (Shannon et al. 2003).

Phylogenetic analyses

The phylogenetic relationships among our samples were determined as part of a larger study (Tepe et al. in prep) with the goal of better resolving hypothesized phylogenetic relationships among *Piper* clades, including *Radula* (Jaramillo *et al.* 2008).

Briefly, a phylogeny was constructed with an 866 base pair region of the *petA-psbJ* gene and a 694 base pair region of the *ITS* gene using the best fit model of evolution (GTR+I+G) across 10,000,000 MCMC iterations in MrBayes 3.2 (Ronquist *et al.* 2012). Forty-eight plant samples from the phylogeny (47.5%; Fig. S1) had chemistry data available for phylogenetic comparative analyses.

We used two complementary analyses to test for phylogenetic signal in the chemical modules (i.e. clusters of co-occurring chemical shifts), λ (Pagel 1999) and *K* (Blomberg *et al.* 2003). λ is a measure of the covariance among traits for tip structure, with values ranging from zero (no phylogenetic signal) to one (strong phylogenetic signal). In contrast, *K* is a ratio of the observed phylogenetic independence of traits to the expected phylogenetic independence under Brownian motion, with $K > 1$ representing signal stronger than expected (e.g., phylogenetic niche conservatism; Losos 2008) and $K < 1$ representing weak or absent signal (Blomberg *et al.* 2003; Revell *et al.* 2008). See Münkemüller *et al.* (2012) for a detailed review of different measures of phylogenetic signal, including a discussion of strengths and weaknesses. For each chemical module, both λ and *K* were estimated with the *phylosig* function in the phytools package (Revell 2012) in R. To test if λ was significantly greater than zero, log-likelihood values from the true λ estimation were compared to λ estimates from a phylogenetic tree constrained to have $\lambda=0$ using a likelihood ratio test. For estimates of *K*, the true estimate was compared to 1,000 tip-shuffling permutations to infer significance. Simulations of both λ (Freckleton *et al.* 2002) and *K* (Blomberg *et al.* 2003) have demonstrated high power (>90%) in detecting phylogenetic signal for phylogenies including more than 40 species.

In addition to testing for phylogenetic signal in univariate chemical traits, we also asked if the evolution of chemical variation was phylogenetically constrained in multivariate space. First, we calculated a multivariate *K* statistic (K_{mult}; Adams 2014) designed for high-dimensional data using the *physignal* function in the geomorph package (Adams & Otárola-Castillo 2013) in R. Additionally, a principal component analysis (PCA) was implemented in R to summarize the chemical variation found within the samples. Both λ and *K* were estimated for the first three principal components (PCs) using the same methods described above for univariate data.

One major concern identified in previous studies is that estimates of phylogenetic signal can be biased if they fail to account for intra-specific measurement error for traits (Ives *et al.* 2007). Despite major concerns regarding the power of Mantel tests for phylogenetic comparative methods (Harmon & Glor 2010), simulations under scenarios of measurement error have found instances where Mantel tests outperform *K* in detecting phylogenetic signal (Hardy & Pavoine 2012). Because we were unable to account for measurement error in our study, we utilized Mantel tests to test for a significant relationship between phylogenetic and chemical distance using the recommendations of Hardy and Pavoine (2012). Two measures of phylogenetic distance were used as predictor variables: first Abouheif's proximity (Abouheif 1999; Pavoine *et al.* 2008) was calculated using the *proxTips* function in the adephylo package (Jombart *et al.* 2010) in R, and second, the square root of patristic distance was calculated using the *cophenetic.phylo* function in the ape package (Paradis *et al.* 2004) in R. Mantel tests were implemented using the *MRM* function in the ecodist package (Goslee & Urban 2007) in R.

We were also interested in asking if there was phylogenetic signal in the diversity of chemical compounds across *Radula*. To estimate signal (λ and *K*) for chemical diversity, we calculated three diversity metrics, Shannon diversity (H), Shannon effective number of chemical compounds (exp(H)), and Simpson effective number of chemical compounds (1/D) (Jost 2006), from the binned NMR spectra data. Each diversity metric was calculated using the entire NMR spectra, as well as using only the downfield or upfield portions of the spectra based on chemical shift (downfield > 5 ppm > upfield). These regions are expected to have different evolutionary histories because downfield resonances are typically associated with deshielded carbon atoms that are indicative of complex chemical structures, while upfield resonances are associated with aliphatic protons that are less structurally diverse (Richards *et al.* in review).

Finally, we examined the relationships among chemical modules to test for potential tradeoffs (characterized by negative correlations) and to illuminate potentially codiversifying chemical traits (characterized by positive correlations) among modules. In addition to calculating simple correlations among modules, we employed a phylogenetic generalized least squares approach (PGLS; Freckleton *et al.* 2002) to calculate phylogenetically controlled correlations among modules using the *pgls* function in the caper package (Orme *et al.* 2013) in R.

RESULTS

Chemical profiles and phylogenetic data were analyzed from 48 specimens (see Table S1 for sampling data and GenBank accession numbers). Based on the scale free topology model fit from the network analyses, 14 chemical modules were selected for

analyses, each comprising a set of at least three related chemical shifts (Fig. 1). By examining the sets of chemical shifts associated with each module, we were able to determine a generic chemical structure characteristic for each module *a posteriori* from the NMR spectra. One network cluster containing three modules had NMR spectra characteristic of lignans, another cluster of three modules corresponded to aliphatics, and a cluster of two modules was characteristic of glycosylated compounds (Fig. 1). In addition, a wide range of chemical compounds corresponded to network clusters composed of a single module, including modules representative of prenyl groups, flavones, and *trans*-cinnamates (Fig. 1).

Most of the modules lacked significant phylogenetic signal for both λ and *K*; however, two chemistry modules displayed weak yet significant phylogenetic signal: Aliphatics 1 (λ = 0.413; *K* = 0.194) and Lignans 3 (λ = 0.242; *K* = 0.177) (Fig. 2; Table S2). In addition, the Flavones module had a significant estimate of signal for λ but not *K* (λ = 0.374; *K* = 0.095) (Fig. 2; Table S2). The first three PCs of chemistry module data explained 27.7, 19.2, and 12.6 percent of the variance in the data, respectively. PC1 had estimates of λ significantly greater than zero (λ = 0.268), but this trend did not hold for estimates of *K* (*K* = 0.117) (Fig. 2; Table S2). The Aliphatics 1 and Lignans 3 modules heavily loaded onto PC1 (Table S3), likely accounting for the detection of phylogenetic signal in PC1. Phylogenetic signal was not detected by either measurement for PC2 or PC3 (Table 1). Furthermore, phylogenetic signal was not recovered using a multivariate *K* statistic (K_{mult} = 0.085; *P* = 0.544). Significant phylogenetic signal was found for only one diversity metric, Simpson effective number of chemical compounds calculated using the entire NMR spectra (*K* = 0.146; *P* = 0.033; λ not significant; Fig. 2; Table S2). While

patristic distance was not a significant predictor of chemical distance (Mantel $R^2 < 0.001$; $P = 0.918$), Abouheif's proximity had a weak yet significant effect (Mantel $R^2 = 0.004$; $P = 0.004$).

Twenty-one of 91 pairwise relationships between chemistry modules were significantly correlated using traditional correlations and PGLS models that correct for phylogenetic independence among species (Table 1). Of these, ten comparisons were negatively correlated and eleven comparisons were positively correlated. Except for the Downfield module, all modules were significantly correlated with at least one other module (Table 1). Of the three chemical modules with phylogenetic signal (Table 1; Fig. 2), Aliphatics 1 and Flavones were both negatively correlated with Lignans 3, but the relationship between Aliphatics 1 and Flavones was non-significant (Table 2).

DISCUSSION

In this study, we were able to rapidly and efficiently generate whole-leaf metabolomic data across a diverse plant lineage to gain a better understanding of the evolution of secondary chemistry across the clade. By focusing on the entire metabolomic profile of species instead of on crude classifications of compound classes, we have been able to elucidate intricate patterns of chemical evolution that are consistent with redundancy (Jones & Firn 1991; Romeo *et al.* 1996), synergy (Dyer *et al*. 2003; Richards *et al*. 2010), and tradeoffs (Johnson *et al.* 2014). In addition, we were able to use the distribution of chemical traits across species to tie metabolomic traits back to chemical structures, allowing for finer resolution of the evolutionary relationships among chemical traits. Few studies have used a phylogenetic comparative approach to

understand patterns of chemical evolution across closely related plants (Johnson *et al.* 2014), leaving many unanswered questions yet to be explored, and future studies would benefit from utilizing metabolomic data to investigate the evolution of chemistry and test long standing hypotheses regarding plant-insect codiversification.

Across 48 species of *Radula*, the evolution of phytochemistry is highly labile as revealed by both univariate (Fig. 2) and multivariate analyses, with only five chemical traits having detectable phylogenetic signal (Aliphatics 1, Lignans 3, Flavones, PC1, $1/D_{(t)}$) (Fig. 2). These results are complementary to a recent study of secondary chemistry in evening primroses (*Oenothera*) (Johnson *et al.* 2014), which also found evidence for high degrees of phytochemical lability among species. In addition, the lack of phylogenetic signal in chemistry across *Radula* suggests that the use of chemical markers as tools for phylogenetic reconstruction, or chemical systematics (Waterman and Gray 1987; Waterman 2007), is likely inappropriate for species rich lineages of plants that are phytochemically diverse. We found mixed effects of phylogenetic signal for lignans, which are ubiquitous compounds with known antiherbivore effects (Jensen et al. 1993; Harmatha & Dinan 2003), as only one of the three lignan modules displayed phylogenetic signal (Fig. 2). Additionally, *Trans*-Cinnamates are key structures found in *Piper* amides that are some of the best characterized secondary compounds with antiherbivore effects in *Piper* (e.g., piplartine and cenocladamide) (Dyer et al. 2003), but this chemical module was highly labile across *Radula* (Fig. 3). Similarly, the Prenyl Groups module did not display significant phylogenetic signal, despite being known precursors of chemical compounds with strong antiherbivore effects in *Piper* species (Jeffrey et al. 2014). Based on the distribution of phylogenetic signal across chemical modules, we are currently

unable to identify a definitive link between the degree of phylogenetic signal and potential antiherbivore effects.

Despite the preponderance of highly evolvable phytochemistry across species in *Radula*, a number of correlations among these traits were discovered (Table 1), also consistent with the results of Johnson *et al.* (2014). This presents an interesting dichotomy of the evolution of chemical variation within versus among species, where perhaps there are a few combinations of chemical traits that confer success, but these combinations are evolutionarily labile from species to species. Of the three lignan modules, Lignans 1 and 3 were both significantly correlated with Lignans 2 using traditional correlations and PGLS (Table 1). Similarly, Alipathatics 1 and 3 were correlated with Aliphatics 2 using both approaches (Table 1). These positive correlations among structurally similar modules likely are a result of these chemical structures being constructed from the same biochemical pathways. On the other hand, aliphatic and lignan modules were typically negatively correlated with one another (Table 1). This result demonstrates the power of this approach to detect meaningful differences in chemical structures across modules because chemical shifts associated with lignans and aliphatics are considered orthogonal; lignans are composed of unshielded and unsaturated hydrogens, while aliphatic hydrogens are shielded and saturated. Overall, many of the correlations among modules are indicative of structural differences; however, these correlations are also consistent with the hypothesis that phytochemical defense is shaped by tradeoffs (Johnson et al. 2014), synergy (Dyer *et al*. 2003; Richards *et al*. 2010), and redundancy (Jones & Firn 1991; Romeo *et al.* 1996) among alternative defensive compounds.

Prior studies have identified a number of issues that could affect estimates of phylogenetic signal in comparative studies (reviewed by Kamilar & Cooper 2013), which are important to discuss within the context of this study. First, the phylogeny of *Radula* is not entirely characterized, so incomplete taxon sampling and unresolved tree structure could have influenced our results. However, we made great effort to sample species from across the entire known phylogeny of *Radula* (Fig. S1) to reduce our sampling bias as much as possible. Furthermore, simulations using incompletely resolved phylogenies (i.e., weak node support) to estimate signal have actually overestimated signal (Davies *et al.* 2012), perhaps making our finding of rampant chemical lability more conservative. In addition, we were unable to quantify the measurement error associated with the chemical traits within species (e.g., Johnson *et al.*, 2014), which can decrease the statistical power for detecting phylogenetic signal (Blomberg *et al.* 2003; Ives *et al.* 2007; Hardy & Pavoine 2012). We attempted to account for this by conducting Mantel tests (which yielded similar results), as they have been demonstrated to be more robust in detecting phylogenetic signal than *K* under some scenarios (Hardy & Pavoine 2012).

It is also possible that environmental effects on our chemical traits potentially could bias estimates of phylogenetic signal and correlations (Ives *et al.* 2007). Despite this, secondary chemistry traits have been consistently shown to have relatively high heritability measurements - nearly twice as large as estimates for other plant functional traits (Geber & Griffen 2003; Johnson *et al.* 2009a) - which somewhat lessens the potential for environmental effects to skew our results. While estimates of chemical heritability are unavailable for the *Radula* species included in this study, we have no reason to expect that *Piper* species deviate from the patterns of high chemical heritability

found in other plant lineages. It is worth noting, however, that while mean heritabilities of chemical compounds are often large, estimates for individual compounds can vary widely across compounds that are from the same branch of a chemical pathway (Caseys *et al.* 2015).

Our results yielded mixed support for the expectation that groups of closely related plants should have similar chemical profiles (Ehrlich and Raven 1964): while individual compounds classes rarely displayed phylogenetic signal (Fig. 2), suites of codiversifying chemical classes were often correlated with one another (Table 1). It is not entirely clear how these dynamics might affect codiversification in plant-insect interactions. In an escape and radiate scenario, phytochemical variation is thought to be largely partitioned among clades, suggesting that herbivores should be able to readily consume all of the plants within a phytochemically similar lineage. In contrast, host shifting between closely related plants might be more difficult in lineages characterized by high chemical lability, which would perhaps select for herbivore specialization (e.g. Forister et al. 2015). Finally, it is important to recognize that the strength of phylogenetic signal depends on phylogenetic scale, as deeper lineages can be readily characterized by their chemical makeup (Fraenkel 1959). A thorough understanding of how phylogenetic scale interacts with the continuum between chemical lability and constraint is needed before plant-insect coevolutionary theory can advance.

ACKNOWLEDGEMENTS

This research was funded by the National Science Foundation (DEB-1145609). We thank Chris Feldman, Beth Leger, and Steve Vander Wall for promiding valuable comments that have improved the text.

REFERENCES

Abouheif, E. (1999). A method for testing the assumption of phylogenetic independence in comparative data. *Evol. Ecol. Res.*, 1, 895-909.

Adams, D. C. (2014). A generalized *K* statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Syst. Biol.*, 63, 685-697.

Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.*, 4, 393-399.

Agrawal, A. A. (2007). Macroevolution of plant defense strategies. *Trends Ecol. Evol.*, 22, 103-109.

Agrawal, A. A. (2011). Current trends in the evolutionary ecology of plant defence. *Funct. Ecol.*, 25, 420-432.

Agrawal, A. A., Fishbein, M., Halitschke, R., Hastings, A. P., Rabosky, D. L. & Rasmann, S. 2009. Evidence of adaptive radiation from a phylogenetic study of plant defenses. *Proc. Natl. Acad. Sci. USA*, 106, 18067-18072.

Berenbaum, M. (1978). Toxicity of a furanocoumarin to armyworms: a case of biosynthetic escape from insect herbivores. *Science*, 201, 532-534.

Berenbaum, M. (1983). Coumarins and caterpillars: a case for evolution. *Evolution*, 37, 163-179.

Blomberg, S. P. & Garland Jr., T. (2002). Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol.*, 15, 899-910.

Blomberg, S. P., Garland Jr., T. & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57, 717-745.

Bowers, M. D. (1983). The role of iridoid glycosides in host-plant specificity of checkerspot butterflies. *J. Chem. Ecol.*, 9, 475-493.

Bowers, M. D. (1984). Iridoid glycosides and host-plant specificity in larvae of the buckeye butterfly, *Junonia coenia* (Nymphalidae). *J. Chem. Ecol.*, 10, 1567-1577.

Caseys, C., Stritt, C., Glauser, G., Blanchard, T. & Lexer, C. 2015. Effects of hybridization and evolutionary constraints on secondary metabolites: the genetic architecture of phenylpropanoids in European *Populus* species. *PLoS ONE*, 10, e0128200.

Davies, T. J., Kraft, N. J. B., Salamin, N. & Wolkovich, E. M. (2012). Incompletely resolved phylogenetic trees inflate estimates of phylogenetic conservatism. *Ecology*, 93, 242-247.

Dyer, L. A. & Palmer, A. D. N. (2004). *Piper: a model genus for studies of phytochemistry, ecology, and evolution*. Kluwer Academic/Plenum, New York.

Dyer, L. A., Dodson, C. D., Beihoffer, J. & Letourneau, D. K. (2001). Trade-offs in antiherbivore defenses in *Piper cenocladum*: ant mutualists versus plant secondary metabolites. *J. Chem. Ecol.*, 27, 581-592.

Dyer, L. A., Dodson, C. D., Stireman, J. O., Tobler, M. A., Smilanich, A. M., Fincher, R. M. *et al.* (2003). Synergistic effects of three *Piper* amides on generalist and specialist herbivores. *J. Chem. Ecol.*, 29, 2499-2514.

Dyer LA, Dodson CD, Richards J (2004a) Piper. A Model Genus for Studies of Evolution, Chemical Ecology, and Trophic Interactions. L. A. Dyer, A. N. Palmer Eds. (Kluwer Academic Publishers, Boston) pp 117-139.

Dyer, L. A., Letourneau, D. K., Dodson, C. D., Tobler, M. A., Stireman, J. O. & Hsu, A. (2004b). Ecological causes and consequences of variation in defensive chemistry of a Neotropical shrub. *Ecology*, 85, 2795-2803.

Ehrlich, P. R. & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 18, 586-608.

Forister, M. L., Novotny, V., Panorska, A. K., Baje, L., Basset, Y., Butterill, P. T. *et al.* (2015). The global distribution of diet breadth in insect herbivores. *Proc. Natl. Acad. Sci. USA*, 112, 442-447.

Fraenkel, G. S. (1959). The raison d'être of secondary plant substances. *Science*, 129, 1466-1470.

Freckleton, R. P., Harvey, P. H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.*, 160, 712-726.

Geber, M. A. & Griffen, L. R. (2003). Inheritance and natural selection on functional traits. *Int. J. Plant Sci.*, 164, S21-S42.

Goslee, S. C. & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.*, 22, 1-19.

Gutierrez, R. M. P., Gonzalez, A. M. N. & Hoyo-Vadillo, C. (2013). Alkaloids from *Piper*: a review of its phytochemistry and pharmacology. *Mini Rev. Med. Chem.*, 13, 163-193.

Hardy, O. J. & Pavoine, S. (2012). Assessing phylogenetic signal with measurement error: a comparison of Mantel tests, Blomberg et al.'s *K*, and phylogenetic distograms. *Evolution*, 66, 2614-2621.

Harmatha, J. & Dinan, L. (2003). Biological activities of lignans and stilbenoids associated with plant-insect chemical interactions. *Phytochem. Rev.*, 2, 321-330.

Harmon, L. J. & Glor, R. E. (2010). Poor statistical performance of the Mantel test in phylogenetic comparative analyses. *Evolution*, 64, 2173-2178.

Horvath, S. (2011). Weighted network analysis. Springer, New York.

Ives, A. R., Midford, P. E. & Garland Jr., T. (2007). Within-species variation and measurement error in phylogenetic comparative methods. *Syst. Biol.*, 56, 252-270.

Janz, N. (2011). Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annu. Rev. Ecol. Evol. Syst.*, 42, 71-89.

Jaramillo, M. A., Callejas, R., Davidson, C., Smith, J. F., Stevens, A. C. & Tepe, E. J. (2008). A phylogeny of the tropical genus *Piper* using ITS and the chloroplast intron *psbJ-petA*. *Syst. Bot.*, 33, 647-660.

Jeffrey, C. S., Leonard, M. D., Glassmire, A. E., Dodson, C. D., Richards, L. A., Kato, M. J., *et al.* (2014) Antiherbivore prenylated benzoic acid derivatives from *Piper kelleyi*. *J. Nat. Prod.*, 77, 148-153.

Jensen, S., Hansen, J. & Boll, P. M. (1993) Lignans and neolignans from Piperaceae. *Phytochemistry*, 33, 523-530.

Johnson, M. T. J., Agrawal, A. A., Maron, J. L. & Salminen, J. P. (2009a). Heritability, covariation and natural selection on 24 traits of common evening primrose (*Oenothera biennis*) from a field experiment. *J. Evol. Biol.*, 22, 1295-1307.

Johnson, M. T. J., Smith, S. D. & Rausher, M. D. (2009b). Plant sex and the evolution of plant defenses against herbivores. *Proc. Natl. Acad. Sci. USA*, 106, 18079-18084.

Johnson, M. T. J., Ives, A. R., Ahern, J. & Salminen, J. P. (2014). Macroevolution of plant defenses against herbivores in the evening primroses. *New Phytol.*, 203, 267-279.

Jombart, T., Balloux, F. & Dray, S. (2010). *adephylo*: new tools for investigating the phylogenetic signal in biological traits. *Bioinformatics*, 26, 1907-1909.

Jones, C. G. & Firn, R. D. (1991). On the evolution of plant secondary chemical diversity. *Phil. Trans. R. Soc. B.*, 333, 273-280.

Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363-375.

Kamilar, J. M. & Cooper, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. *Phil. Trans. R. Soc. B.*, 368, 20120341.

Karinho-Betancourt, E., Agrawal, A. A., Halitschke, R. & Nunez-Farfan, J. 2015. Phylogenetic correlations among chemical and physical plant defenses change with ontogeny. *New Phytol.*, 206, 796-806.

Kato, M. J. & Furlan, M. (2007) Chemistry and evolution of the Piperaceae. Pure Appl. Chem. 79: 529-538.

Krishnan, P., Kruger, N. J. & Ratcliffe, R. G. (2005). Metabolite fingerprinting and profiling in plants using NMR. *J. Exp. Bot.*, 56, 255-265.

Kursar, T. A., Dexter, K. G., Lokvam, J., Pennington, T., Richardson, J. E., Weber, M. G. *et al.* (2009). The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proc. Natl. Acad. Sci. USA*, 106, 18073-18078.

Langfelder, P. & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC *Bioinform.*, 9, 559.

Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.*, 11, 995-1007.

Macel, M., van Dam, N. M. & Keurentjes, J. J. B. (2010). Metabolomics: the chemistry between ecology and genetics. *Mol. Ecol. Res.*, 10, 583-593.

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., *et al.* (2012). How to measure and test phylogenetic signal. *Methods Ecol. Evol.*, 3, 743-756.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. *et al.* (2013). caper: comparative analyses of phylogenetics and evolution in R. R package version 0.5.2. http://CRAN.R-project.org/package=caper

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401, 877-884.

Paradis, E., Claude, J. & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289-290.

Parmar VS, Jain, S. C., Bisht, K. S., Jain, R., Taneja, P., Jha, A., *et al.* (1997). Phytochemistry of the genus *Piper*. Phytochemistry, 46, 597- 673.

Pavoine, S., Ollier, S., Pontier, D. & Chessel, D. (2008). Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theor. Popul. Biol.*, 73, 79-91.

R Core Team. (2013). R: language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 URL http://www.R-project.org/

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.*, 3, 217-223.

Revell, L. J., Harmon, L. J. & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Syst. Biol.*, 57, 591-601.

Richards, L. A., Dyer, L. A., Smilanich, A. M., & Dodson, C. D. (2010). Synergistic effects of amides from two *Piper* species on generalist and specialist herbivores. *J. Chem. Ecol.*, 36, 1105-1113.

Richards, L.A., Lampert, E.C., Bowers, M.D., Dodson, C.D., Smilanich, A.M. & Dyer, L.A. (2012) Synergistic effects of iridoid glycosides on the survival, development, and immune response of a specialist caterpillar (*Junonia coenia* Nymphalidae). *J. Chem. Ecol.*, 38, 1276-1284.

Richards, L. A., Dyer, L. A., Forister, M. L., Smilanich, A. M., Dodson, C. D., Leonard, M. D., *et al.* (in review). Phytochemical diversity drives diversity of tropical plant-insect communities. *Proc. Natl. Acad. Sci. USA*

Romeo, J. T., Saunders, J. A. & Barbosa, P. (1996). *Phytochemical diversity and redundancy in ecological interactions*. Plenum Press, New York.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., *et al.* (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.*, 61, 539-542.

Rønsted, N., Symonds, M. R. E., Birkholm, T., Christensen, S. B., Meerow, A. W., Molander, M. *et al.* (2012). Can phylogeny predict chemical diversity and potential medicinal activity of plants? A case study of amaryllidaceae. *BMC Evol. Biol.*, 12, 182.

Saslis-Lagoudakis, C. H., Klitgaard, B. B., Forest, F., Francis, L., Savolainen, V., Williamson, E. M. *et al.* (2011). The use of phylogeny to interpret cross-cultural patterns in plant use and guide medicinal plant discovery: an example from *Pterocarpus* (Leguminosae). *PLoS ONE*, 6, e22275.

Saslis-Lagoudakis, C. H., Savolainen, V., Williamson, E. M., Forest, F., Wagstaff, S. J., Baral, S. R. *et al.* (2012). Phylogenies reveal predictive power of traditional medicine in bioprospecting. *Proc. Natl. Acad. Sci. USA*, 109, 15835-15840.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498-2504.

Smilanich, A. M., Dyer, L. A., Chambers, J. Q. & Bowers, M. D. (2009). Immunological cost of chemical defence and the evolution of herbivore diet breadth. *Ecol. Lett.*, 12, 612-621.

Tepe, E. J., Rodríguez-Castañeda, G., Glassmire, A. E. & Dyer, L. A. (2014). *Piper kelleyi*, a hotspot of ecological interactions and a new species from Ecuador and Peru. *PhytoKeys*, 34, 19-32.

Thompson, J. N. (1989). Concepts of coevolution. *Trends Ecol. Evol.*, 4, 179-183.

Thompson, J. N. & Pellmyr, O. (1991). Evolution of oviposition behavior and host preference in Lepidoptera. *Annu. Rev. Entomol.*, 36, 65-89.

Waterman, P. G. (2007). The current status of chemical systematics. *Phytochemistry*, 68, 2896-2903.

Waterman, P. G. & Gray, A. I. (1987). Chemical systematics. *Nat. Prod. Rep.*, 4, 175-203.

Yessoufou, K., Daru, B. H. & Muasya, A. M. (2015). Phylogenetic exploration of commonly used medicinal plants in South Africa. *Mol. Ecol. Res.*, 15, 405-413.

Table 1. Traditional correlations (lower triangle) and phylogenetic generalized least squares models (upper triangle; *β*-coefficients from models reported) were used to test for relationships among chemical traits. Significant tests (α = 0.05) are denoted with bold text.

	Aliphatics 1	Glycosylated Phenols	Glycosylated Compounds	Unknown Aromatics	Prenyl Groups	*Trans*-Cinnamate	Ligands 1	Ligands 3	Ligands 2	Aliphatics 3	Aliphatics 2	Downfield	Flavones	Background
Aliphatics 1	-	**-4.579**	-1.323	-2.45	-0.259	**-5.235**[a]	**-12.574**[a]	**-2.264**[a]	**-3.975**[a]	0.333	**0.704**[a]	0.525	2.25	0.056
Glycosylated Phenols	-0.203	-	0.025	0.274	-0.044	**0.731**[a]	**1.185**[a]	**0.126**	0.068	-0.013	-0.007	**-0.082**	0.012	-0.013
Glycosylated Compounds	-0.09	0.235	-	-0.67	**-0.535**[a]	-0.254	-0.401	-0.137	0.134	**-0.128**[a]	**-0.124**[a]	0.046	-0.022	-0.069
Unknown Aromatics	**-0.511**	0.187	-0.196	-	**0.287**[a]	**0.400**[a]	0.463	0.068	-0.117	0.005	0.002	0.018	**-0.295**	-0.03
Prenyl Groups	**-0.516**	-0.035	**-0.405**[a]	**0.454**[a]	-	0.006	-0.474	0.151	-0.386	0.1	**0.089**	-0.002	-0.009	**0.104**[a]
Trans-Cinnamate	**-0.477**[a]	**0.587**[a]	-0.017	**0.501**[a]	0.113	-	**1.555**[a]	0.075	0.041	**-0.040**[a]	-0.033	-0.023	-0.13	-0.037
Ligands 1	**-0.446**[a]	**0.392**[a]	0	**0.313**	-0.032	**0.653**[a]	-	0.033	**0.173**[a]	-0.018	**-0.019**[a]	-0.005	-0.037	**-0.024**
Ligands 3	**-0.601**[a]	0.125	-0.173	**0.333**	**0.46**	0.235	0.014	-	**1.582**[a]	0.049	0.01	-0.077	**-0.589**[a]	0.032
Ligands 2	**-0.530**[a]	0.167	0.056	0.097	0.023	0.283	**0.366**[a]	**0.665**[a]	-	-0.021	-0.02	0.066	-0.151	-0.042
Aliphatics 3	0.17	-0.226	**-0.579**[a]	-0.103	0.238	**-0.361**[a]	**-0.392**	0.031	-0.197	-	**0.806**[a]	-0.622	1.318	0.386
Alipatics 2	**0.420**[a]	-0.072	**-0.511**[a]	-0.133	0.029	**-0.286**	**-0.349**[a]	-0.233	**-0.34**	**0.799**[a]	-	0.013	1.509	0.282
Downfield	0.042	-0.187	0.048	0.013	0.003	-0.114	-0.044	-0.071	0.085	-0.055	0.152	-	**-0.956**	-0.151
Flavones	0.014	0.106	0.033	-0.144	0.163	-0.013	0.123	**-0.337**[a]	-0.221	0.157	0.212	-0.21	-	**0.091**[a]
Background	-0.088	-0.03	-0.263	-0.182	**0.399**[a]	-0.103	-0.236	0.081	-0.172	**0.385**	**0.354**	-0.068	**0.387**[a]	-

[a] Significant pairwise relationships among chemical traits using both traditional correlations and phylogenetic generalized least squares models.

Figure Legends

Fig. 1. A network displays the relationships among each binned chemical shift. The number within each node represents a unique chemical shift (ppm) and each edge depicts pairs of chemical shifts that often co-occur in NMR spectra. Chemical shifts with the same color belong to the same chemical module. Chemical shifts from the Background module were not included in network analyses for ease of visualization, as these chemical shifts are highly interconnected with all modules and greatly outnumber other chemical shifts.

Fig. 2. The distribution of phylogenetic signal across 26 chemical traits is displayed for two complementary estimates, λ and *K* (panels A and B, respectively). Estimates for the 14 chemical modules are shown in black, the 3 PCs are shown in white, and the 9 diversity metrics are shown in gray. Chemical diversity was calculated using three metrics [Shannon diversity (H); Shannon effective number of chemical compounds (exp(H)); Simpson effective number of chemical compounds (1/D)] using the entire NMR spectra (total), as well as for subsets including only the upfield and downfield ranges of the spectra (down > 5ppm > up). Asterisks denote significant estimates of phylogenetic signal ($\alpha = 0.05$).

Fig. 3. The distribution of chemical variation in five traits (Aliphatics 1, Lignans 3, Flavones, Prenyl Groups, *Trans*-Cinnamates) across the phylogeny of *Radula* is

displayed. Module colors match those from Fig. 1. Each species value for a chemical trait was divided by the maximum trait value to standardize values for easy visualization. Significant estimates of both λ and *K* were found for Aliphatics 1 and Lignans 3, but estimates for Flavones were only significant for λ (Fig. 2; Table S2). See Fig. S1 for additional information about the phylogenetic tree structure, including node support.

Figure 1

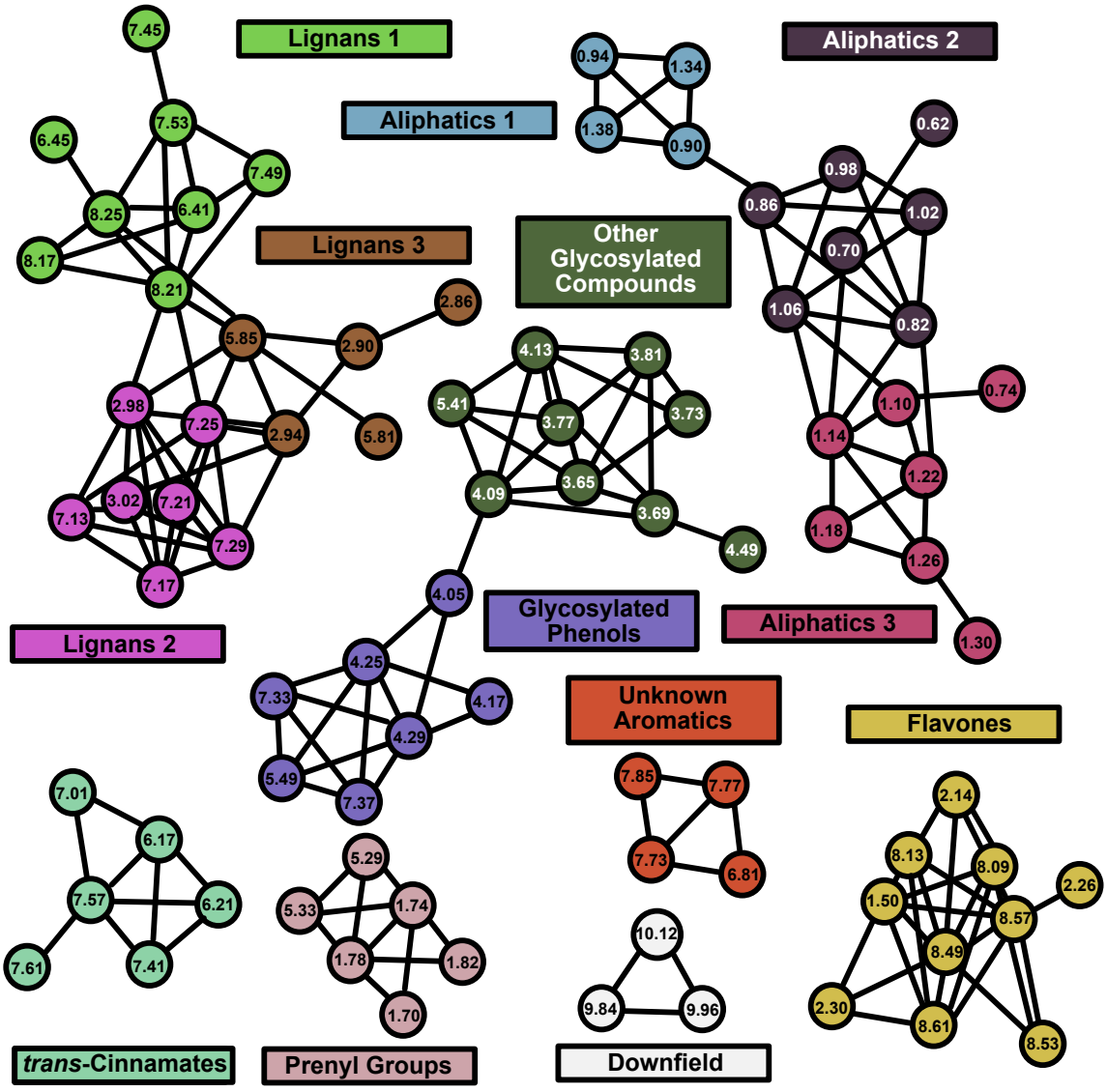

Lignans 1
Aliphatics 1
Aliphatics 2
Lignans 3
Other Glycosylated Compounds
Glycosylated Phenols
Aliphatics 3
Lignans 2
Unknown Aromatics
Flavones
trans-Cinnamates
Prenyl Groups
Downfield

Figure 2

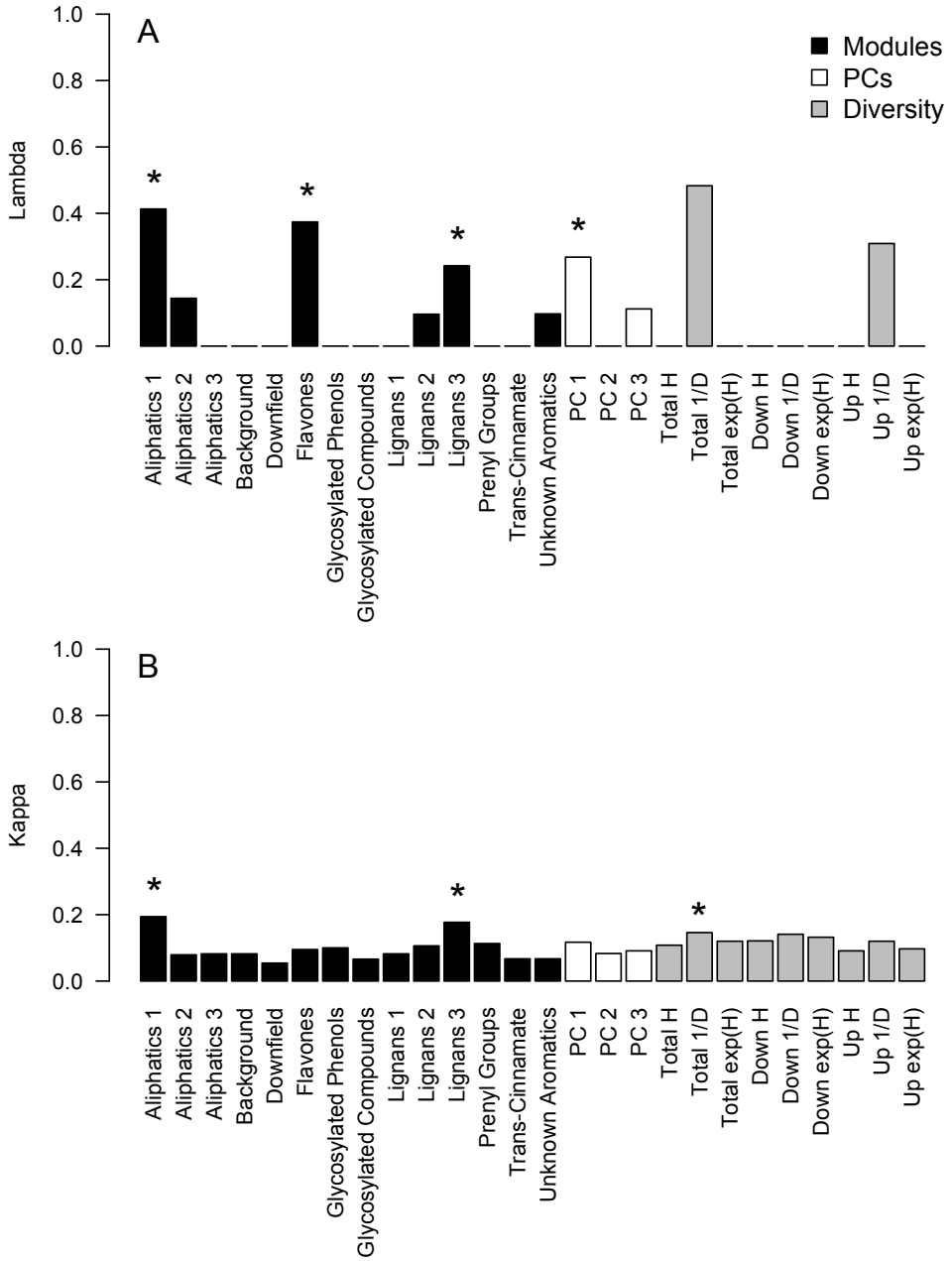
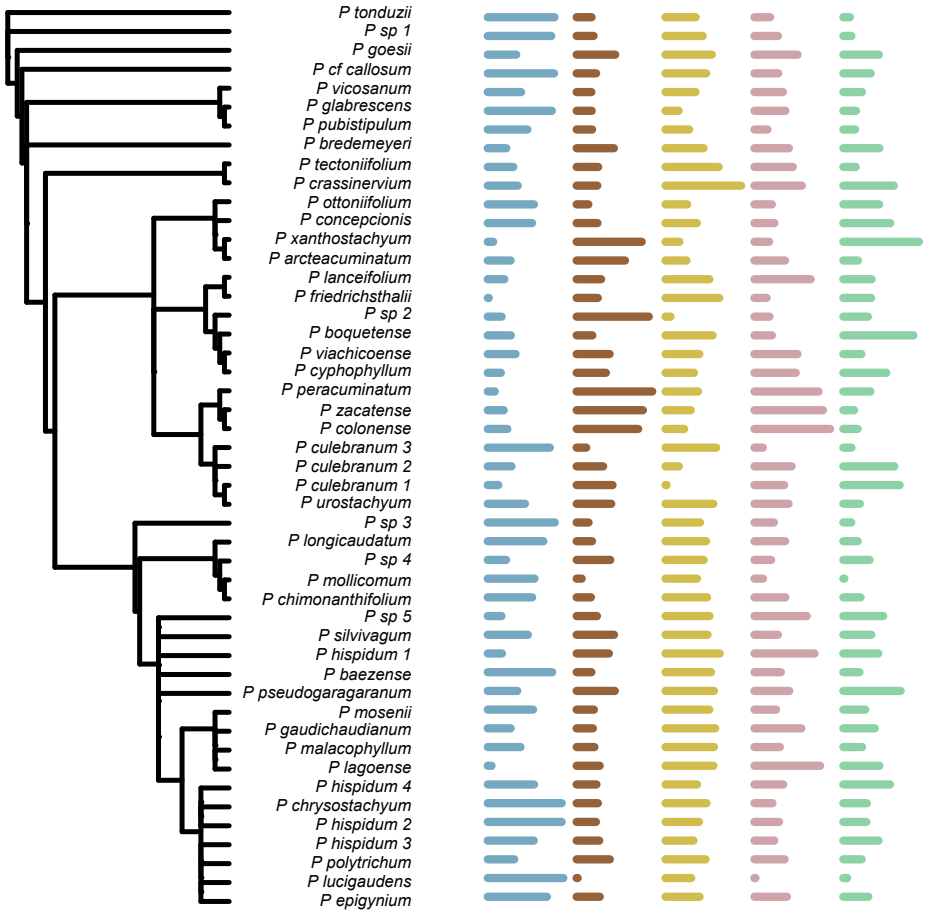

A
Modules
PCs
Diversity
Lambda
1.0
0.8
0.6
0.4
0.2
0.0
Aliphatics 1
Aliphatics 2
Aliphatics 3
Background
Downfield
Flavones
Glycosylated Phenols
Glycosylated Compounds
Lignans 1
Lignans 2
Lignans 3
Prenyl Groups
Trans-Cinnamate
Unknown Aromatics
PC 1
PC 2
PC 3
Total H
Total 1/D
Total exp(H)
Down H
Down 1/D
Down exp(H)
Up H
Up 1/D
Up exp(H)
B
Kappa

Figure 3


P tonduzii
P sp 1
P goesii
P cf callosum
P vicosanum
P glabrescens
P pubistipulum
P bredemeyeri
P tectoniifolium
P crassinervium
P ottoniifolium
P concepcionis
P xanthostachyum
P arcteacuminatum
P lanceifolium
P friedrichsthalii
P sp 2
P boquetense
P viachicoense
P cyphophyllum
P peracuminatum
P zacatense
P colonense
P culebranum 3
P culebranum 2
P culebranum 1
P urostachyum
P sp 3
P longicaudatum
P sp 4
P mollicomum
P chimonanthifolium
P sp 5
P silvivagum
P hispidum 1
P baezense
P pseudogaragaranum
P mosenii
P gaudichaudianum
P malacophyllum
P lagoense
P hispidum 4
P chrysostachyum
P hispidum 2
P hispidum 3
P polytrichum
P lucigaudens
P epigynium

SUPPLEMENTARY MATERIAL

Chapter 2 ~ Phylogenetic signal and the evolution of secondary chemistry in *Radula* (Piperaceae): persistent trait correlations despite rampant trait lability

Joshua P. Jahner, Eric J. Tepe, Lora A. Richards, Matthew L. Forister, Lee A. Dyer, Angela M. Smilanich, Craig D. Dodson, Christopher S. Jeffrey

Table S1. Collection information for each *Radula* specimen included in this study, including GenBank accession numbers for the *ITS*, *petA*, and *psbJ* genes. Codes for duplicated species names match those from Fig. 3 in the main text.

Species	Code	Country	Year	Collector	Voucher (herbarium)	*ITS*	*petA-psbJ*
P. arcteacuminatum Trel.	3534	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3534* (CR)	KT007500	KT030286
P. baezense Trel.	YY1	Ecuador	2012	A.E. Glassmire	*A.E. Glassmire* (QCNE)	KT007501	KT030287
P. boquetense Yunck.	3991	Panama	2013	E.J. Tepe	*E.J. Tepe 3991* (PMA)	KT007502	KT030288
P. bredemeyeri J.Jacq.	3495	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3495* (CR)	KT007503	KT030289
P. cf. *callosum* Ruiz & Pav.	3490	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3490* (CR)	KT007543	KT030291
P. chimonanthifolium Kunth	K-1960	Brazil	2014	M. Kato	*M. Kato K-1960* (SPF)	KT007504	KT030290
P. chrysostachyum C.DC.	3482	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3482* (CR)	KT007505	KT030292
P. colonense C.DC.	3502	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3502* (CR)	KT007506	KT030293
P. concepcionis Trel.	3990	Panama	2013	E.J. Tepe	*E.J. Tepe 3990* (PMA)	KT007507	KT030294
P. crassinervium Kunth	K-1954	Brazil	2014	M. Kato	*M. Kato K-1954* (SPF)	KT007508	KT030295
P. culebranum C.DC. (1)	3413	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3413* (CR)	KT007509	KT030296
P. culebranum C.DC. (2)	3527	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3527* (CR)	KT007510	KT030297
P. culebranum C.DC. (3)	4005	Panama	2013	E.J. Tepe	*E.J. Tepe 4005* (PMA)	KT030329	KT030298
P. cyphophyllum C.DC.	3444	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3444* (CR)	KT007511	KT030299
P. epigynium C.DC.	3457	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3457* (CR)	KT007512	KT030300
P. friedrichsthalii C.DC.	131	Costa Rica	2000	E.J. Tepe	*E.J. Tepe 131* (CR)	KT007513	KT030301
P. gaudichaudianum Kunth	K-1949	Brazil	2014	M. Kato	*M. Kato K-1949* (SPF)	KT007514	KT030302
P. glabrescens (Miq.)C.DC.	4022	Panama	2013	E.J. Tepe	*E.J. Tepe 4022* (PMA)	KT007515	KT030303
P. goesii Yunck.	K-1964	Brazil	2014	M. Kato	*M. Kato K-1964* (SPF)	KT007544	KT030304
P. hispidum Sw. (1)	3430	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3430* (CR)	KT007516	KT030305
P. hispidum Sw. (2)	3496	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3496* (CR)	KT007517	—
P. hispidum Sw. (3)	3509	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3509* (CR)	KT007518	KT030306
P. hispidum Sw. (4)	3537	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3537 (CR)*	KT007519	—
P. lagoense C.DC.	K-1944	Brazil	2014	M. Kato	*M. Kato K-1944* (SPF)	KT007520	KT030307
P. lanceifolium Kunth	184	Costa Rica	2000	E.J. Tepe	*E.J. Tepe 184* (CR)	KT007521	KT030308
P. longicaudatum Trel. & Yunck.	1578	Ecuador	2006	E.J. Tepe	*E.J. Tepe 1578* (QCNE)	KT007522	KT030309
P. lucigaudens C.DC.	3993	Panama	2013	E.J. Tepe	*E.J. Tepe 3993* (PMA)	KT007523	KT030310
P. malacophyllum (C.Presl.) C.DC.	K-1945	Brazil	2014	M. Kato	*M. Kato K-1945* (SPF)	KT007524	KT030311
P. mollicomum Kunth	K-1942	Brazil	2014	M. Kato	*M. Kato K-1942* (SPF)	KT007525	KT030312
P. mosenii C.DC.	K-1948	Brazil	2014	M. Kato	*M. Kato 1948* (SPF)	KT007526	KT030313
P. ottoniifolium C.DC.	4011	Panama	2013	E.J. Tepe	*E.J. Tepe 4011* (PMA)	KT007527	KT030314
P. peracuminatum C.DC.	3433	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3433* (CR)	KT007528	KT030315
P. polytrichum C.DC.	3965	Panama	2013	E.J. Tepe	*E.J. Tepe 3965* (CR)	KT007529	KT030316
P. pseudogaragaranum Trel.	4063	Panama	2013	E.J. Tepe	*E.J. Tepe 4063* (PMA)	KT007530	KT030317
P. pubistipulum C.DC.	4027	Panama	2013	E.J. Tepe	*E.J. Tepe 4027* (PMA)	KT007531	KT030318
P. silvivagum C.DC.	3523	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3523* (CR)	KT007532	KT030320
P. sp. (1)	3536	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3536* (CR)	KT007545	KT030327
P. sp. (2)	4056	Panama	2013	E.J. Tepe	*E.J. Tepe 4056* (PMA)	KT007534	KT030321
P. sp. (3)	4996	Ecuador	2011	J. Homeier	*J. Homeier 4996* (LOJA)	KT007533	KT030319
P. sp. (4)	K-1978	Brazil	2014	M. Kato	*M. Kato K-1978* (SPF)	KT007536	KT030323
P. sp. (5)	K-1983	Brazil	2014	M. Kato	*M. Kato K-1983* (SPF)	KT007537	KT030322
P. tectoniifolium Kunth	K-1958	Brazil	2014	M. Kato	*M. Kato K-1958* (SPF)	KT007535	—
P. tonduzii C.DC.	3532	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3532* (CR)	KT007546	KT030328
P. urostachyum Hemsl.	3535	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3535* (CR)	KT007538	—
P. viachicoense Yunck.	4051	Panama	2013	E.J. Tepe	*E.J. Tepe 4051* (PMA)	KT007539	—
P. vicosanum Yunck.	K-1966	Brazil	2014	M. Kato	*M. Kato K-1966* (SPF)	KT007540	KT030324
P. xanthostachyum C.DC.	3542	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3542* (CR)	KT007541	KT030325
P. zacatense C.DC.	3438	Costa Rica	2012	E.J. Tepe	*E.J. Tepe 3438* (CR)	KT007542	KT030326

Table S2. Summary statistics from two tests of phylogenetic signal, λ (Pagel 1999) and *K* (Blomberg *et al*. 2003), for 26 chemistry traits. The three diversity metrics (Shannon diversity [H]; Shannon effective number of chemical compounds [exp(H)]; Simpson effective number of chemical compounds [1/D]) were calculated using the entire NMR spectra (t), as well as using only the upfield (u) and downfield (d) portions of the NMR spectra. For each estimate of λ, the log-likelihood (LL) of the true estimate was compared to the log-likelihood of a constrained model (where λ=0) using a likelihood ratio test to determine if λ is significantly larger than zero. For each estimate of *K*, significance is inferred by comparing the true estimate to 1,000 tree tip permutations. Significant signal is denoted with bold text and asterisks (α = 0.05).

Trait	λ	LL	LL (λ=0)	*P*	*K*	*P*
Aliphatics 1	0.413	45.4	42.8	**0.024***	0.194	**0.004***
Aliphatics 2	0.144	44.8	44.2	0.273	0.079	0.655
Aliphatics 3	0	41	41	1	0.082	0.586
Background	0	38.4	38.4	1	0.082	0.579
Downfield	0	19	19	1	0.054	0.937
Flavones	0.374	-45.5	-47.8	**0.031***	0.095	0.362
Glycosylated Compounds	0	-25.9	-25.9	1	0.066	0.847
Glycosylated Phenols	0	-57	-57	1	0.1	0.309
Lignans 1	0	-92.1	-92.1	1	0.082	0.59
Lignans 2	0.096	-59.9	-60.2	0.462	0.106	0.267
Lignans 3	0.242	-26	-28.3	**0.033***	0.177	**0.016***
Prenyl Groups	0	-27.4	-27.4	1	0.113	0.177
Trans-Cinnamates	0	-48.7	-48.7	1	0.067	0.831
Unknown Aromatics	0.097	-48.5	-49.1	0.271	0.067	0.828
PC 1	0.268	-98.2	-100.6	**0.030***	0.117	0.129
PC 2	0	-91.9	-91.9	1	0.083	0.581
PC 3	0.112	-81.3	-81.7	0.384	0.091	0.419
$H_{(t)}$	0	-7.8	-7.8	1	0.108	0.208
$exp(H)_{(t)}$	0	-163.3	-163.3	1	0.12	0.101
$1/D_{(t)}$	0.483	-191.4	-191	1	0.146	**0.033***
$H_{(d)}$	0	-5.1	-5.1	1	0.121	0.119
$exp(H)_{(d)}$	0	-144.1	-144.1	1	0.132	0.093
$1/D_{(d)}$	0	-162.6	-162.6	1	0.141	0.135
$H_{(u)}$	0	-6.8	-6.8	1	0.091	0.449
$exp(H)_{(d)}$	0	-155.5	-155.5	1	0.097	0.325
$1/D_{(d)}$	0.309	-179	-178.6	1	0.12	0.096

Table S3. Loadings for the PCA summarizing chemical modules in multivariate space.

Module	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
Aliphatics 1	0.385	0.254	0	0.229	-0.135	0.114	0	-0.209	-0.485	-0.251	0.132	0	0.253	0.518
Aliphatics 2	0.35	-0.258	-0.121	0.322	-0.235	-0.208	-0.148	0.209	0.116	-0.102	0	0	-0.686	0.163
Aliphatics 3	0.306	-0.383	0	0.115	-0.263	0	0	0.353	0.304	0.227	-0.232	0	0.568	0.147
Background	0.158	-0.343	-0.213	-0.403	0	-0.239	-0.151	-0.544	0.318	-0.166	0.341	0	0.13	0.102
Downfield	0	0	0.301	0.269	0.296	-0.795	-0.166	0	-0.113	0	0	-0.152	0.167	-0.122
Flavones	0.105	0	-0.554	-0.302	0.131	-0.22	0.277	0.381	-0.318	-0.349	0	-0.183	0	-0.171
Glycosylated Compounds	-0.124	0.429	0	-0.384	0.107	-0.144	-0.366	0.363	0.287	0	0	0	0	0.513
Glycosylated Phenols	-0.242	0	-0.402	0	-0.366	0	-0.575	0	-0.277	0.313	0.203	0	0.14	-0.24
Lignans 1	-0.344	0	-0.297	0.245	0	-0.199	0.491	0	0.106	0.371	0.277	-0.303	0	0.35
Lignans 2	-0.321	0	0.278	-0.131	-0.463	-0.289	0.266	0.165	-0.135	-0.244	0.137	0.525	0.104	0.139
Lignans 3	-0.272	-0.32	0.35	-0.199	-0.278	0	-0.148	0	-0.188	-0.169	0	-0.684	-0.138	0
Prenyl Groups	0	-0.488	0	-0.162	0.395	0	0	0	-0.427	0.363	-0.171	0.27	0	0.371
Trans-Cinnamates	-0.378	0	-0.299	0.269	0	0	-0.1	-0.3	0.133	-0.339	-0.653	0	0	0.144
Unknown Aromatics	-0.279	-0.244	0	0.366	0.391	0.219	-0.176	0.291	0.138	-0.394	0.448	0	0.178	0

Figure S1

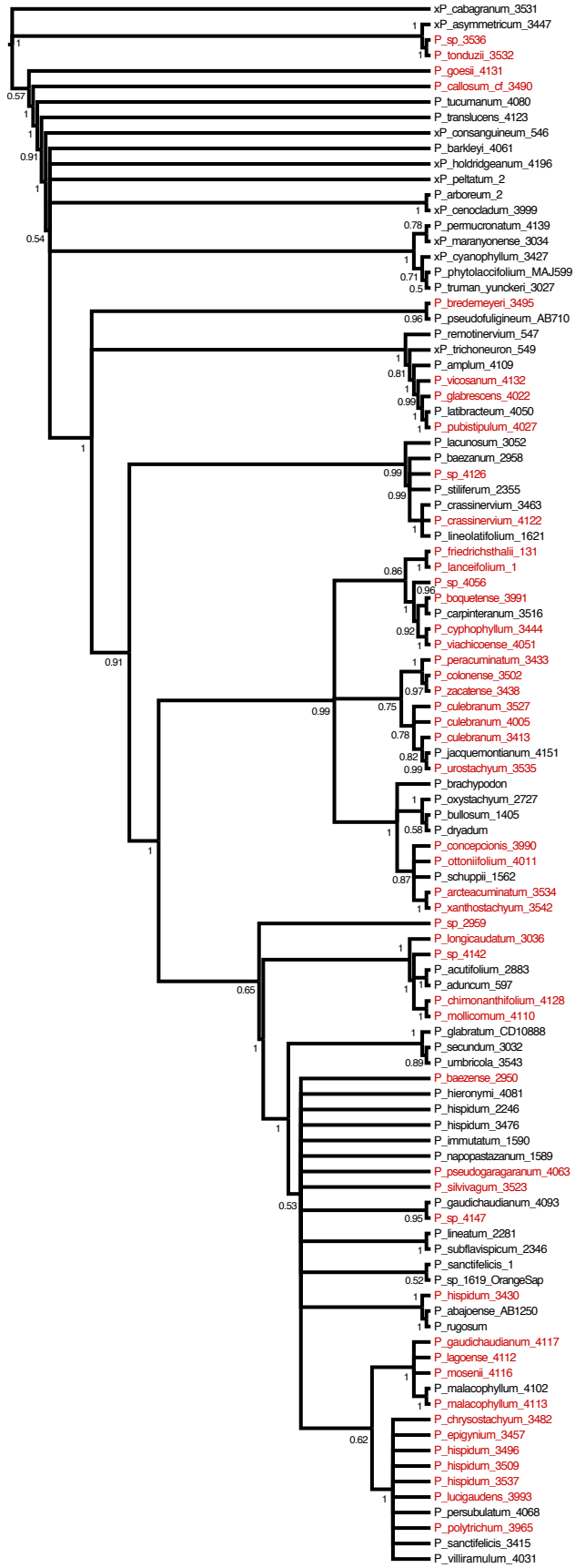

xP_cabagranum_3531
xP_asymmetricum_3447
P_sp_3536
P_tonduzii_3532
P_goesii_4131
P_callosum_cf_3490
P_tucumanum_4080
P_translucens_4123
xP_consanguineum_546
P_barkleyi_4061
xP_holdridgeanum_4196
xP_peltatum_2
P_arboreum_2
xP_cenocladum_3999
P_permucronatum_4139
xP_maranyonense_3034
xP_cyanophyllum_3427
P_phytolaccifolium_MAJ599
P_truman_yunckeri_3027
P_bredemeyeri_3495
P_pseudofuligineum_AB710
P_remotinervium_547
xP_trichoneuron_549
P_amplum_4109
P_vicosanum_4132
P_glabrescens_4022
P_latibracteum_4050
P_pubistipulum_4027
P_lacunosum_3052
P_baezanum_2958
P_sp_4126
P_stiliferum_2355
P_crassinervium_3463
P_crassinervium_4122
P_lineolatifolium_1621
P_friedrichsthalii_131
P_lanceifolium_1
P_sp_4056
P_boquetense_3991
P_carpinteranum_3516
P_cyphophyllum_3444
P_viachicoense_4051
P_peracuminatum_3433
P_colonense_3502
P_zacatense_3438
P_culebranum_3527
P_culebranum_4005
P_culebranum_3413
P_jacquemontianum_4151
P_urostachyum_3535
P_brachypodon
P_oxystachyum_2727
P_bullosum_1405
P_dryadum
P_concepcionis_3990
P_ottoniifolium_4011
P_schuppii_1562
P_arcteacuminatum_3534
P_xanthostachyum_3542
P_sp_2959
P_longicaudatum_3036
P_sp_4142
P_acutifolium_2883
P_aduncum_597
P_chimonanthifolium_4128
P_mollicomum_4110
P_glabratum_CD10888
P_secundum_3032
P_umbricola_3543
P_baezense_2950
P_hieronymi_4081
P_hispidum_2246
P_hispidum_3476
P_immutatum_1590
P_napopastazanum_1589
P_pseudogaragaranum_4063
P_silvivagum_3523
P_gaudichaudianum_4093
P_sp_4147
P_lineatum_2281
P_subflavispicum_2346
P_sanctifelicis_1
P_sp_1619_OrangeSap
P_hispidum_3430
P_abajoense_AB1250
P_rugosum
P_gaudichaudianum_4117
P_lagoense_4112
P_mosenii_4116
P_malacophyllum_4102
P_malacophyllum_4113
P_chrysostachyum_3482
P_epigynium_3457
P_hispidum_3496
P_hispidum_3509
P_hispidum_3537
P_lucigaudens_3993
P_persubulatum_4068
P_polytrichum_3965
P_sanctifelicis_3415
P_villiramulum_4031

Fig. S1. A phylogenetic tree depicts the evolutionary relationships from all Radula species for which genetic data has been accumulated (Tepe et al. in prep). Individuals that were included in this study's analyses are labeled in red text. Posterior probabilities are listed for each node to indicate support.

Figure S2

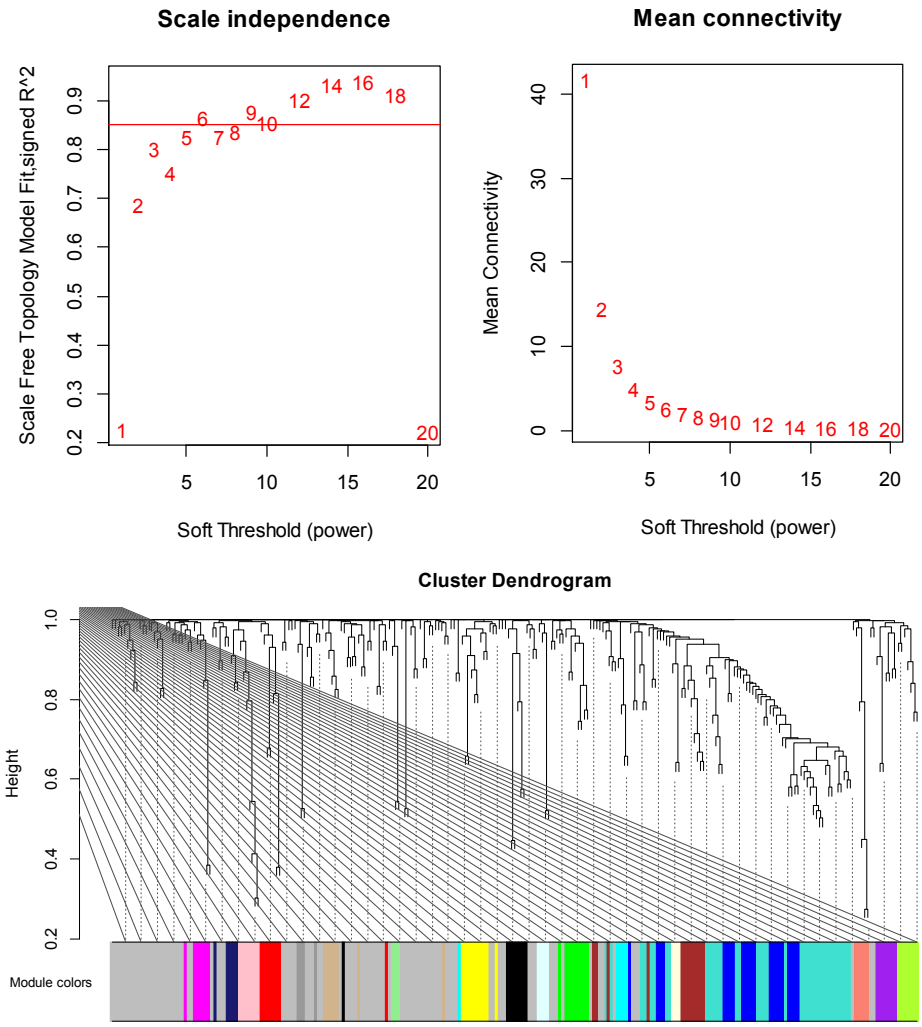


Figure S2. Analyses used in module construction. A) Analysis to determine the soft threshold to satisfy scale free topology. The red line indicates the power functions where $R^2 > 0.85$. B) The mean connectivity of the network at varying power transformations. C) The cluster dendrogram of chemical shifts with dissimilarity based on topological overlap and module construction below. The gray module is considered the spectral baseline.

Chapter 3 ~ Use of exotic hosts by Lepidoptera: widespread species colonize more novel hosts

Joshua P. Jahner,[1,2] Melvin M. Bonilla,[1] Kevin J. Badik,[3] Arthur M. Shapiro,[4] and Matthew L. Forister[1]

[1]*Program in Ecology, Evolution, and Conservation Biology, Department of Biology, University of Nevada, Reno, NV 89557, U.S.A.*

[2]*Email: jpjahner@msn.com*

[3]*Program in Ecology, Evolution, and Conservation Biology, Department of Natural Resources and Environmental Science, University of Nevada, Reno, NV, 89557, U.S.A.*

[4]*Center for Population Biology, University of California, Davis, CA 95616, U.S.A.*

Abstract: The study of host shifts by herbivorous insects has played an important role in evolutionary biology, contributing to research in coevolution, ecological speciation, and adaptive radiation. As invasive plants become more abundant in many ecosystems, the potential for exotic host use by native insects increases. Graves and Shapiro (2003) have documented exotic host use by 34% of Californian butterflies, suggesting that the plants and butterflies of California might be an important model system for the colonization and utilization of novel resources. In this study, we analyze relationships among geographic range, native diet breadth, and the use of exotic hosts for Californian butterflies and skippers (Lepidoptera). Geographic range and, to a lesser extent, native diet breadth are significant predictors of exotic host use, with positive correlations found both before and after phylogenetic correction. These results give insight into the process of insect host range evolution, as geographically widespread generalists have an apparently greater tendency to utilize novel hosts than geographically constrained specialists. Increasing occurrences of exotic host use are expected and those species not capable of shifting to non-native hosts are likely to have higher vulnerability to extirpation and extinction in the future.

KEY WORDS: Diet breadth, geographic range, independent contrasts, oscillation hypothesis

INTRODUCTION

Interactions between herbivorous insects and exotic plants have become model systems in a number of areas of evolutionary biology, from the study of diversification to rapid adaptation associated with anthropogenic environmental change (Drès and Mallet 2002; Matsubayashi et al. 2009). In a few well-studied examples, hosts shifts involving exotic plants have resulted in speciation (e.g. *Rhagoletis* flies shifting from hawthorn to apple trees; Feder 1998; Feder and Forbes 2008), the evolution of feeding morphology (e.g. *Leptocoris* bugs adapting to balloon vine; Carroll 2008; Carroll et al. 2005a), and shifts in oviposition behavior (e.g. *Euphydryas* butterflies and exotic *Plantago* hosts; Singer et al. 1993; Singer et al. 2008). A shift to an exotic host has also been associated with an escape from natural enemies (*Papilio machaon aliaska* caterpillars on novel Asteraceae hosts; Murphy 2004), as well as an increased reliance on a protection mutualism (*Lycaeides melissa* caterpillars with ants on alfalfa; Forister et al. 2011). Outside of these detailed studies, we have much to learn about the frequency with which native insects utilize exotic plants, the potential for host shifts, and the conditions which either constrain or promote the colonization of exotic hosts (Thompson 1998; Strauss et al. 2006).

An understanding of conditions associated with insect host shifts is also important in light of the oscillation hypothesis (Janz and Nylin 2008), which has been recently proposed as an explanation for herbivorous insect diversification through alternating periods of dietary specialization and generalization within lineages. According to the oscillation hypothesis, an expansion of host breadth in a previously specialized taxon is associated with an increase in geographic range, which facilitates subsequent local adaptation to new hosts, specialization, and ultimately diversification. Mechanisms of

specialization have been studied, including trade-offs in feeding efficiency and limitations in neural capacity (Bernays 2001; Fox and Morrow 1981; Janz and Nylin 1997; Nylin 1988). What has been less investigated is the relationship between host breadth and geographic range, though the recent finding of a positive correlation between host range and geographic range in the butterfly subfamily Nymphalinae provides support for the oscillation hypothesis (Janz and Nylin 2008; Slove and Janz 2011).

The butterflies, skippers, and plants of California provide an excellent system to study the colonization of novel plants, as at least 34% of Californian butterfly species have been reported as utilizing non-native plant species (Graves and Shapiro 2003). In this study, we use exotic host plant records from Graves and Shapiro (2003) in combination with native host data and published geographic ranges for 70 species (five families) of butterflies and skippers to investigate relationships among exotic host use, native host range, and geographic range. In particular, we ask if geographic range size and native diet breadth are significant predictors of exotic host use using both raw and phylogenetically-corrected data.

MATERIALS AND METHODS

Study Organisms

Seventy species of butterflies (Papilionoidea) and skippers (Hesperioidea) reported by Graves and Shapiro (2003) as having exploited native plant species were included in this study. These species include representatives of five Lepidopteran families: Hesperiidae, Lycaenidae, Nymphalidae, Papilionidae, and Pieridae.

Data Collection

Graves and Shapiro (2003) conducted a literature review and interviewed Californian lepidopterists to produce a list of known exotic host plants for Lepidoptera found in California. In the list of exotic hosts reported by Graves and Shapiro (2003), each host record for every butterfly and skipper was assigned a confidence level ranging from "high" to "unlikely". Exotic hosts were only included in our study if the records were reported with a "high" or "moderate" level of confidence. Graves and Shapiro (2003) classified "high" confidence as host records that are "well documented" in California and "moderate confidence" as host records that "seemed reasonable given known foodplants of the butterfly and distributions of both butterfly and plant" in California. Plant hosts with lower confidence levels were excluded from this study because Graves and Shapiro (2003) deemed these records "unlikely" and their inclusion could have overestimated exotic diet breadth in our study. All exotic host records are of caterpillar feeding, as "oviposition only" records were not included in this study. Note that the records reported by Graves and Shapiro (2003) are associations between Lepidopterans and exotic hosts that do not necessarily imply successful utilization by a butterfly or skipper, since herbivorous insects have been known to oviposit on plants that are ultimately lethal to their offspring (Keeler and Chew 2008). Thus, we are examining associations between native butterflies and exotic hosts without knowing which plants are viable hosts. Nevertheless, the association is of inherent interest and represents the first stage of contact between native insects and exotic plants, even if the association is detrimental or transient in some unknown proportion of cases. Butterfly and skipper species with zero exotic host plant records were excluded from this study (11 species) because we were only interested in species reported to have utilized exotics.

Native diet breadth was recorded as the number of native plant genera used by caterpillars, based on observations of one of us (A.M.S.), and from field guides (Garth and Tilden 1986; Scott 1986); see Appendix S1 for native and exotic diet breadths. Body size was originally included as a potential predictor of exotic host use due to previous findings documenting adult body size as a predictor of butterfly extinctions (Koh et al. 2004); however, associations with body size were weak and non-significant and are not discussed further.

Geographic range size was calculated using county records obtained from Opler et al. (2010); see Appendix S1. Specifically, we calculated the total square kilometers for occupied counties for each species. County records were used in this study due to their availability for the contiguous 48 states (Opler et al. 2010). Range sizes from the continental United States were used instead of Californian range sizes under the assumption that butterfly exotic host shifts occur throughout the United States, and Californian shifts are just one subset. Furthermore, using county records for the 48 states avoided artifacts associated with range records restricted to California. For example, a species with a continent-wide distribution might only occur in a small part of California; including county records from the 48 states allowed us to properly record such a species as widespread (records are also available for Canada and Mexico, though the very large counties in parts of these countries might exaggerate geographic range). This methodology is supported by the results of Hawkins and Porter (2003), who found that relationships in butterfly richness are conserved when sampling across geographic scales.

Phylogenetic data and independent contrasts

As no comprehensive phylogenetic hypothesis exists for North American Lepidoptera, mitochondrial DNA sequences from a portion of the cytochrome oxidase subunit I region (COI) were used to construct a phylogeny to be used in the generation of independent contrasts. COI sequences were selected for use because they were available for most species included in this study. Sequences were preferentially chosen from GenBank in order to have the most complete COI sequences possible; see Appendix S1 for accession numbers and for criteria in choosing accessions. Sequences were aligned with Clustal as implemented in Sequencher 4.10.1. Being a single, rapidly evolving marker, COI data is not equally informative at all taxonomic levels (Forister et al. 2008); therefore, a phylogeny was constrained in BEAUti 1.6.1 by creating monophyletic taxon sets based on previously reported phylogenetic hypotheses (Drummond and Rambaut 2007). Two separate trees were constrained for analysis due to recent hypotheses of the family level relationships of Lepidoptera; see Appendix S1 for methods used in constructing constraints. Using the constrained trees and COI data, Bayesian searches were performed in BEAST 1.6.1 using a GTR + Invariate + Gamma substitution model (Drummond and Rambaut 2007); see Appendix S1 for a figure of one phylogeny used. Phylogenetic independent contrasts of range size, native host breadth, and exotic host breadth were calculated using the APE package in R 2.11.1 (Felsenstein 1985; Paradis et al. 2004; R Development Core Team 2010).

Analyses

To address our central question regarding relationships among geographic range, native host breadth, and the number of exotic hosts colonized, we conducted analyses in two phases: first using raw data (prior to phylogenetic correction), then using data

following phylogenetic independent contrasts. For both sets of analyses (before and after phylogenetic correction), we used simple linear models in which the response variable was the number of exotic hosts colonized and the predictor variables were geographic range and native host breadth (the number of native genera). To investigate colonization dynamics that might be affected by the diversity of hosts used, analyses were performed using (as the response variable) log-transformations of four taxonomic indices: number of exotic species, number of exotic genera, exotic species multiplied by genera, and exotic genera multiplied by families. For the models involving un-corrected data, we also included family as a categorical variable, and investigated interactions between family and both of the predictors (family is not a relevant category for the data following phylogenetic correction). Following phylogenetic corrections, data were highly leptokurtotic and were normalized with Johnson Su transformations (Slifker and Shapiro 1980). All linear models were calculated in JMP 8.0.

To account for the relatedness of exotic plants with native hosts, another set of analyses was done with exotic plants categorized as congeneric, confamilial (but not congeneric), or non-confamilial relative to the herbivore's native host range based on the methods of Connor et al. (1980). For example, the report of *Lycaeides melissa* using the plant *Medicago sativa* (Fabaceae) was categorized as confamilial because *L. melissa* does not have native *Medicago* hosts, but does have native hosts in Fabaceae. All plant taxonomy was determined using the USDA plants database (http://plants.usda.gov/). Each exotic host record was then ranked by relatedness to native hosts (congeneric = 1; confamilial = 2; non-confamilial = 3), and the taxonomic isolation of exotic hosts for each butterfly or skipper was calculated as the mean rank of all reported exotic hosts

from Graves and Shapiro (2003). A linear regression was used to investigate the relationship between geographic range and the taxonomic isolation of exotic hosts utilized.

RESULTS

Geographic range and native host plant breadth are both significantly associated with the number of exotics colonized, and this is true both with the raw data before phylogenetic correction and with the data subsequent to correction (raw data is shown in Fig. 1). For the uncorrected data, 36% of the variation was explained in the number of exotic host species used by the 70 included butterfly and skipper species ($R^2 = 0.36$; $F_{6,63} = 5.96$; $P < 0.0001$). Geographic range and native host breadth were both significant predictors, with standardized beta coefficients of 0.45 ($F_{1,63} = 16.74$; $P = 0.0001$) and 0.30 ($F_{1,63} = 8.27$; $P = 0.0055$) respectively (beta coefficients are coefficients from linear models that have been calculated in units of standard deviations to facilitate comparisons among predictor variables measured on different scales, such as geographic range and native host range (Zar 2010, pg. 433)). Interactions between family and both geographic range and native host breadth were not significant and were dropped from the model. The main effect of family was similarly not significant ($F_{4,63} = 0.78$; $P = 0.54$). The two predictor variables of geographic range and native host breadth were also not significantly correlated with each other (Spearman's rank correlation: $\rho = 0.06$; $P = 0.65$). Finally, we found no relationship between geographic range and the taxonomic isolation of exotic hosts relative to native host breadth ($R^2 = 0.0003$; $F_{1,68} = 0.02$; $P = 0.88$) (this relationship was not tested using phylogenetic independent contrasts). In other words, while more widespread species are more likely to colonize more new hosts, these hosts

are not more likely to represent major host shifts (e.g. to new families). Similar results (i.e. associations between geographic range, native host use and exotic host use) using uncorrected data were found across analyses for all taxonomic indices of exotic host use; see Appendix S2 for results from linear models for all indices (as explained above, other indices included the number of exotic genera, the number of exotic species multiplied by the number of exotic genera, and the number of exotic genera multiplied by the number of exotic families).

The phylogenetically corrected linear model including geographic range and native host breadth accounted for 37% of the variation in exotic host use ($R^2 = 0.37$; $F_{2,66} = 18.97$; $P < 0.0001$). Geographic range and native host breadth remained significant predictors of exotic host use after phylogenetic correction. Geographic range had a standard beta coefficient of 0.51 ($F_{1,66} = 26.31$; $P < 0.0001$) and native host range had a standard beta coefficient of 0.27 ($F_{1,66} = 7.44$; $P = 0.0082$). As with the analyses using raw data, the effect of geographic range on exotic colonization is greater than the effect of native host breadth (judged by beta coefficients). Qualitatively similar results were found with phylogenetically corrected data from all phylogenies; see Appendix S2 for results from all constraints using number of species and number of genera to quantify exotic host use.

DISCUSSION

These results reveal a potentially simple facet of the evolution of host range expansion: geographically widespread species colonize more exotic hosts. This pattern holds across taxonomic families, which suggests a simple biological mechanism by which greater exposure to exotic hosts provides greater opportunities for successful

colonization. We also find that species with a broader native host range have colonized more exotic hosts, which is perhaps less surprising than the association between geographic range and the number of exotic hosts colonized. These findings support earlier results of Carroll et al. (2005b) who found that the most polyphagous and geographically widespread Australian soapberry bug colonized the most exotic plants. This suggests that relationships among geographic range, native host range, and exotic host use might be important in many phytophagous insect groups. When comparing the relative importance of geographic range and native host breadth for predicting exotic colonizations, we find that the former (geographic range) is a stronger predictor (standardized beta coefficients discussed above; also see Fig. 1). Furthermore, it is important to note that geographic range and native diet breadth are not correlated with one another ($\rho = 0.06$).

The result that generalists are more likely to shift to exotic hosts than specialists also complements previous research by Parker and others, who reported that generalist herbivores have a higher preference for and can confer more damage to invasive species than specialists (Parker and Hay 2005; Parker et al. 2006). The phytochemicals in plants that butterflies use as cues to pick oviposition sites might be partly responsible for this pattern. Herbivores are more likely to colonize plant species if they have already encountered the plants and associated phytochemicals or if the hosts are utilized by closely related species (Janz et al. 2001; Liu et al. 2005); therefore, herbivores with small ranges are less likely to colonize exotic plants with unfamiliar suites of phytochemicals that are not present in their geographic range (Cappuccino and Arnason 2006). In contrast, herbivores with large geographic ranges are potentially more likely to shift to

exotic hosts because they have a greater probability of having previously encountered a wide array of plants and phytochemicals (of course, this would not necessarily be true of species with large ranges but small distances travelled by individual insects).

In addition to being generally consistent with the oscillation hypothesis of herbivorous insect diversification (Janz and Nylin 2008), our results including exotic host species provide a valuable complement to the positive correlation between butterfly geographic range and diet breadth reported by Slove and Janz (2011). We assume that the geographic ranges of the butterflies in our study have been relatively static compared to the recent invasion of exotic plants in California. Thus the finding that geographically widespread butterflies colonize more exotic hosts than more localized species suggests a direction of causality in which wide geographic range can influence host range expansion. Future research in this area could profitably focus on the mechanistic factors that promote the expansion of geographic range and host breadth using butterflies and skippers in California as model systems. Potential promoters of specialist host expansion include mutations for generalist behavior and phenotypic plasticity on novel hosts followed by genetic accommodation (Nylin and Janz 2009; Weingartner et al. 2006).

Up to two-thirds of the 2000 plus endemic plants found in California are projected to undergo range reductions greater than 80% of their current range over the next 100 years (Loarie et al. 2008). A reduction in the abundance and availability of native plants to herbivores combined with an accelerating rate of invasion by exotic species into California (Cohen and Carlton 1998) creates strong evolutionary and ecological pressures selecting for exotic host use in herbivorous insects. The effects of these forces should be most apparent in urban areas where native plants experience elevated extinction rates

(Hahs et al. 2009). In fact, some urban Californian butterflies are now entirely dependent on exotic host plants for larval development (Shapiro 2002). For species unable to shift to exotics, this extirpation of native hosts might exacerbate the combined negative effects of habitat loss and climate change (Forister et al. 2010). However, herbivorous insects are not equally prone to extinction, as generalists had lower extinction rates during the Cretaceous mass extinction than specialists (Labandeira et al. 2002). The same pattern can also be found in the current mass extinction, with coextinction proneness positively correlated with the degree of specialization between an organism and its hosts (Dunn et al. 2009). As native plants rapidly disappear, widespread generalists are more likely to colonize exotic hosts and are less likely to experience coextinction with their traditional host plants than localized specialists.

ACKNOWLEDGEMENTS

We thank the Biology Department at the University of Nevada, Reno, for support. For assistance in obtaining Lepidopteran county records we thank P. A. Opler, K. Lotts, and T. Naberhaus of Butterflies and Moths of North America, for assistance with Lepidopteran phylogeny we thank A. D. Warren, and for assistance with genetic analyses we thank J. S. Wilson. An earlier version was greatly improved with comments from C. D. Jiggins, two anonymous reviewers, and the University of Nevada, Reno EECB peer review group.

LITERATURE CITED

Bernays, E. A. 2001. Neural limitations in phytophagous insects: implications for diet breadth and evolution of host affiliation. Annu. Rev. Entomol. 46:703-727.

Cappuccino, N., and J. T. Arnason. 2006. Novel chemistry of invasive exotic plants. Biol. Lett. 2:189-193.

Carroll, S. P. 2008. Facing change: forms and foundations of contemporary adaptation to biotic invasions. Mol. Ecol. 17:361-372.

Carroll, S. P., J. E. Loye, H. Dingle, M. Mathieson, T. R. Famula, and M. P. Zalucki. 2005a. And the beak shall inherit – evolution in response to invasion. Ecol. Lett. 8:944-951.

Carroll, S. P., J. E. Loye, H. Dingle, M. Mathieson, and M. P. Zalucki. 2005b. Ecology of *Leptocoris* Hahn (Hemiptera: Rhopalidae) soapberry bugs in Australia. Aust. J. Entomol. 44:344-353.

Cohen, A. N., and J. T. Carlton. 1998. Accelerating invasion rate in a highly invaded estuary. Science 279:555-558.

Connor, E. F., S. H. Faeth, D. Simberloff, and P. A. Opler. 1980. Taxonomic isolation and the accumulation of herbivorous insects: a comparison of introduced and native trees. Ecol. Entomol. 5:205-211.

Drès, M., and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Phil. Trans. R. Soc. Lond. B. 357:471-492.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Dunn, R. R., N. C. Harris, R. K. Colwell, L. P. Koh, and N. S. Sodhi. 2009. The sixth mass coextinction: are most endangered species parasites and mutualists? Proc. R. Soc. Lond. B. 276:3037-3045.

Feder, J. L. 1998. The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom about speciation? Pp. 130-144 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford Univ. Press, Oxford, UK.

Feder, J. L., and A. A. Forbes. 2008. Host fruit-odor discrimination and sympatric host-race formation. Pp. 101-116 *in* K. J. Tilmon, ed. Specialization, speciation and radiation: the evolutionary biology of herbivorous insects. Univ. California Press, Berkeley, CA.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1-15.

Forister, M. L., Z. Gompert, C. C. Nice, G. W. Forister, and J. A. Fordyce. 2011. Ant association facilitates the evolution of diet breadth in a lycaenid butterfly. Proc. R. Soc. Lond. B. doi:10.1098/rspb.2010.1959.

Forister, M. L., A. C. McCall, N. J. Sanders, J. A. Fordyce, J. H. Thorne, J. O'Brien, D. P. Waetjen, and A. M. Shapiro. 2010. Compounded effects of climate change and habitat alteration shift patterns of butterfly diversity. Proc. Natl. Acad. Sci. USA 107:2088-2092.

Forister, M. L., C. C. Nice, J. A. Fordyce, Z. Gompert, and A. M. Shapiro. 2008. Considering evolutionary processes in the use of single-locus genetic data for conservation, with examples from the Lepidoptera. J. Insect. Conserv. 12:37-51.

Fox, L. R., and P. A. Morrow. 1981. Specialization: species property or local phenomenon? Science 211:887-893.

Garth, J. S., and J. W. Tilden. 1986. California butterflies. Univ. California Press, Berkeley, CA.

Graves, S. D., and A. M. Shapiro. 2003. Exotics as host plants of the California butterfly fauna. Biol. Conserv. 110:413-433.

Hahs, A. K., M. J. McDonnell, M. A. McCarthy, P. A. Vesk, R. T. Corlett, B. A. Norton, S. E. Clemants, R. P. Duncan, K. Thompson, M. W. Schwartz, and S. G. Williams. 2009. A global synthesis of plant extinction rates in urban areas. Ecol. Lett. 12:1165-1173.

Hawkins, B. A., and E. E. Porter. 2003. Does herbivore diversity depend on plant diversity? The case of California butterflies. Am. Nat. 161:40-49.

Janz, N., and S. Nylin. 1997. The role of female search behaviour in determining host plant range in plant feeding insects: a test of the information processing hypothesis. Proc. R. Soc. Lond. B. 264:701-707.

Janz, N., and S. Nylin. 2008. The oscillation hypothesis of host-plant range and speciation. Pp. 203-215 *in* K. J. Tilmon, ed. Specialization, speciation and radiation: the evolutionary biology of herbivorous insects. Univ. California Press, Berkeley, CA.

Janz, N., K. Nyblom, and S. Nylin. 2001. Evolutionary dynamics of host plant specialization: a case study of the tribe Nymphalini. Evolution 55:783-796.

Keeler, M. S., and F. S. Chew. 2008. Escaping an evolutionary trap: preference and performance of a native insect on an exotic host. Oecologia 156:559-568.

Koh, L. P., N. S. Sodhi, and B. W. Brook. 2004. Ecological correlates of extinction proneness in tropical butterflies. Conserv. Biol. 18:1571-1578.

Labandeira, C. C., K. R. Johnson, and P. Wilf. 2002. Impact of the terminal Cretaceous event on plant-insect associations. Proc. Natl. Acad. Sci. USA 99:2061-2066.

Liu, S.-S., Y.-H. Li, Y.-Q. Liu, and M. P. Zalucki. 2005. Experience-induced preference for oviposition repellents derived from a non-host plant by a specialist herbivore. Ecol. Lett. 8:722-729.

Loarie, S. R., B. Carter, K. Hayhoe, R. Moe, C. A. Knight, D. D. Ackerly. 2008. Climate change and the future of California's endemic flora. PLoS ONE 3:e2502.

Matsubayashi, K. W., I. Ohshima, and P. Nosil. 2009. Ecological speciation in phytophagous insects. Entomol. Exp. Appl. 134:1-27.

Murphy, S. M. 2004. Enemy-free space maintains swallowtail butterfly host shift. Proc. Natl. Acad. Sci. USA 101:18048-18052.

Nylin, S. 1988. Host plant specialization and seasonality in a polyphagous butterfly, *Polygonia c-album* (Nymphalidae). Oikos 53:381-386.

Nylin, S., and N. Janz. 2009. Butterfly host plant range: an example of plasticity as a promoter of speciation? Evol. Ecol. 23:137-146.

Opler, P.A., K. Lotts, and T. Naberhaus, coordinators. 2010. Butterflies and moths of North America. Bozeman, MT: Big Sky Institute. http://www.butterfliesandmoths.org/ (Version 09/28/2010)

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.

Parker, J. D., and M. E. Hay. 2005. Biotic resistance to plant invasions? Native herbivores prefer non-native plants. Ecol. Lett. 8:959-967.

Parker, J. D., D. E. Burkepile, and M. E. Hay. 2006. Opposing effects of native and exotic herbivores on plant invasions. Science 311:1459-1461.

R Development Core Team. 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Scott, J. A. 1986. The butterflies of North America. Stanford Univ. Press, Stanford, CA.

Shapiro, A. M. 2002. The Californian urban butterfly fauna is dependent on alien plants. Divers. Distrib. 8:31-40.

Singer, M. C., C. D. Thomas, and C. Parmesan. 1993. Rapid human-induced evolution of insect host associations. Nature 366:681-683.

Singer, M. C., B. Wee, S. Hawkins, and M. Butcher. 2008. Rapid natural and anthropogenic diet evolution: three examples from checkerspot butterflies. Pp. 311-324 *in* K. J. Tilmon, ed. The evolutionary biology of herbivorous insects: specialization, speciation, and radiation. Univ. California Press, Berkeley, CA.

Slifker, J. F., and S. S. Shapiro. 1980. The Johnson system: selection and parameter estimation. Technometrics 22:239-246.

Slove, J., and N. Janz. 2011. The relationship between diet breadth and geographic range size in the butterfly subfamily Nymphalinae – a study of global scale. PLoS ONE 6:e16057.

Strauss, S. Y., J. A. Lau, and S. P. Carroll. 2006. Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? Ecol. Lett. 9:357-374.

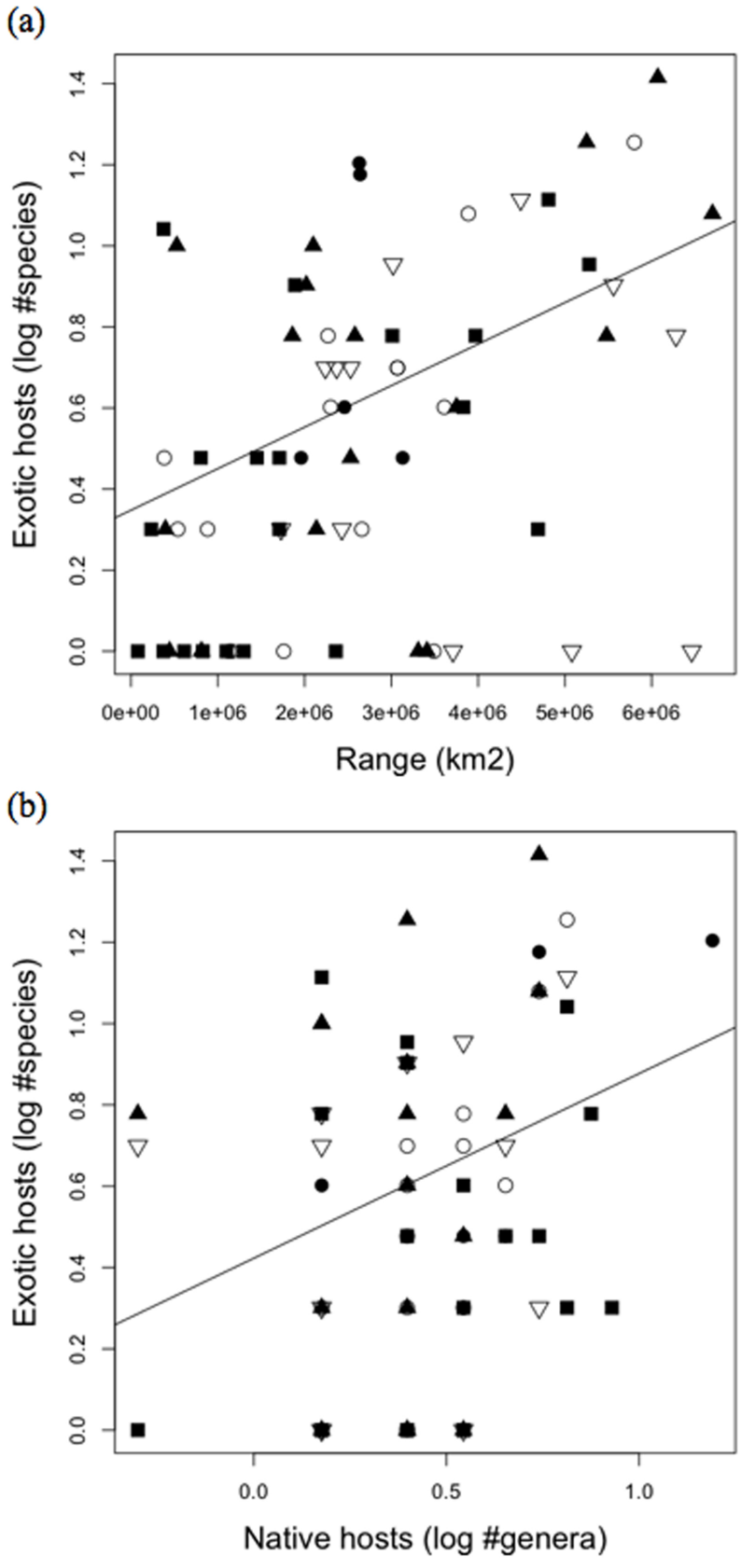
Thompson, J. N. 1998. Rapid evolution as an ecological process. Trends Ecol. Evol. 13:329-332.

Weingartner, E., N. Wahlberg, and S. Nylin. 2006. Dynamics of host plant use and species diversity in *Polygonia* butterflies (Nymphalidae). J. Evol. Biol. 19:483-491.

Zar, J. H. 2010. Biostatistical analysis. 5th ed. Prentice Hall, Upper Saddle River, NJ.

Figure Legend

Fig. 1. Predictors of exotic host use for the 70 species of Lepidoptera included in this study. Data are presented (without phylogenetic correction) for the relationships between the number of exotic host genera used and (a) geographic range ($y = 0.348 + 1.024e^{-7}(x)$) and (b) native diet breadth ($y = 0.423 + 0.454(x)$). Symbols indicate taxonomic families as follows: filled triangle, Pieridae; filled circle, Papilionidae; filled square, Hesperiidae; open triangle, Nymphalidae; open circle, Lycaenidae.


(a)
Exotic hosts (log #species)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0e+00
1e+06
2e+06
3e+06
4e+06
5e+06
6e+06
Range (km2)
(b)
Exotic hosts (log #species)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.0
0.5
1.0
Native hosts (log #genera)

Supplementary Material

Use of exotic hosts by Lepidoptera:

widespread species colonize more

novel hosts

Joshua P. Jahner, Melvin M. Bonilla, Kevin J. Badik, Arthur M. Shapiro, & Matthew L. Forister

Contents:

Appendix S1. *Criteria for selecting sequences, methods for the construction of constraint trees, GenBank accession numbers of COI sequences used, raw ecological data used for phylogenetic independent contrasts, and a figure of one phylogeny used.*

Appendix S2. *Criteria for selecting sequences, methods for the construction of constraint trees, GenBank accession numbers of COI sequences used, raw ecological data used for phylogenetic independent contrasts, and a figure of one phylogeny used.*

Appendix S1. *Criteria for selecting sequences, methods for the construction of constraint trees, GenBank accession numbers of COI sequences used, raw ecological data used for phylogenetic independent contrasts, and a figure of one phylogeny used.*

A complete mitochondrial COI sequence was the first criterion in selecting accession numbers for included species (by complete we refer to all of the 658bp fragment that is commonly sequenced in DNA "barcoding" studies). If a complete COI sequence was unavailable for a species, a congeneric species with a complete sequence was selected. If a congeneric sequence was unavailable, recently published Lepidopteran phylogenies were used to determine the closest related genus with a complete, available sequence (see below). For the following species, a closely related genus could not be determined from a recently published phylogeny so closely related genera were selected based on recommendations from A. D. Warren: *Brephidium exilis*, *Hemiargus ceraunus*, and *Hemiargus isola*; see Table S1 for substitutions. If a species had a complete sequence available while another species within the same genus did not, the same sequence was used for both species (6 species).

To ensure that known relationships within Rhopalocera (butterflies and skippers) were included in the tree, a backbone was constructed using previously published phylogenies. Three separate trees were constructed for analyses due to recent publications describing new hypotheses of the family level relationships in the Lepidoptera. The first tree was constructed with a family constraint of Hesperiidae + (Papilionidae + (Pieridae + (Nymphalidae + Lycaenidae))) (Scoble 1986; Wahlberg et al. 2005a), the second tree with a family constraint of Papilionidae + (Hesperiidae + (Pieridae + (Nymphalidae +

Lycaenidae))) (Mutanen et al. 2010; Regier et al. 2009), and the third tree with no constraints between any species in the tree. The moths *Spodoptera frugiperda* (Noctuidae) and *Rhodoneura aurata* (Thyrididae) were used as outgroups, with the relative positions of Noctuidae and Thyrididae to Rhopalocera based on the phylogeny of Mutanen et al. (2010). Outgroups were not included in subsequent phylogenetic independent analyses.

Known relationships within Hesperiidae families and tribes were added to the backbone structure of the first two constraint trees using data from Warren et al. (2008). Similarly Chew and Watt's (2006) and Zakharov et al.'s (2004) phylogenies were used to create structure within the Pieridae and Papilionidae respectively. Wahlberg et al.'s (2003) phylogeny was used for the subfamily structure within Nymphalidae, and genus level relationships within the subfamily Nymphalinae (Nymphalidae) were determined using Wahlberg et al. (2005b). Species in Lycaenidae were grouped by tribe using Pierce et al. (2002) and subfamily relationships were based on the hypothesis of Pohl et al. (2009). A polytomy existed within the genus *Lycaena* (Lycaenidae) due to the availability of only one adequate COI sequence for the five included species in the genus. Since polytomies violate requirements for calculating independent contrasts in APE, a topology for *Lycaena* was constructed using Pratt and Wright's (2002) phylogeny of North American coppers.

Table S1. Geographic ranges, host plant range (native and exotic), and accession numbers for sequences used in the study. Ecological data are not listed for *Rhodoneura aurata* and *Spodoptera frugiperda* because they were outgroups in phylogenetic trees and not included in analyses.

Species in Tree	GenBank Species Used	GenBank Accession	Geographic Range (km^2)	Native Genera	Exotic Species	Exotic Genera	Exotic Families
Aglais milberti	*Aglais urticae*	HQ003952	3712903	1	1	1	1
Agraulis vanillae	*Agraulis vanillae*	GU333740	2235231	0	5	1	1
Anthocharis lanceolata	*Anthocharis cardamines*	HQ003961	445076	2	1	6	1
Anthocharis sara	*Anthocharis cardamines*	HQ003961	527521	1	10	6	1
Atalopedes campestris	*Atalopedes campestris*	GU089672	3973089	7	6	3	2
Brephidium exilis	*Zizeeria knysna*	AY556972	2299429	2	4	1	1
Calpodes ethlius	*Calpodes ethlius*	GU149470	612568	0	1	1	1
Colias alexandra	*Colias philodice*	GU089776	2136723	1	2	5	1
Colias eurytheme	*Colias eurytheme*	GU089775	6701999	5	12	4	1
Colias philodice	*Colias philodice*	GU089776	5484106	0	6	1	1
Copaeodes aurantiaca	*Ancyloxypha numitor*	GU089635	1296154	3	1	5	1
Cupido comyntas	*Cupido comyntas*	GU089786	3887564	5	12	2	2
Danaus gilippus	*Danaus gilippus*	DQ071865	2431890	1	2	1	1
Danaus plexippus	*Danaus plexippus*	DQ018954	6463445	1	1	2	1
Epargyreus clarus	*Epargyreus clarus*	GU089841	4690602	3	2	1	1
Erynnis funeralis	*Erynnis tristis*	GU155965	1707483	5	3	1	1
Erynnis tristis	*Erynnis tristis*	GU155965	827632	1	1	6	1
Euchloe ausonides	*Euchloe ausonia*	HQ004469	2096686	1	10	2	1
Euchloe hyantis	*Euchloe ausonia*	HQ004469	396868	2	2	4	3
Euphydryas chalcedona	*Euphydryas chalcedona*	AF187752	2531086	4	5	1	1
Euphydryas editha	*Euphydryas editha*	AF187765	1729953	5	2	1	1
Eurema nicippe	*Eurema nicippe*	GU089560	3405383	2	1	1	1
Glaucopsyche lygdamus	*Glaucopsyche lygdamus*	FJ808850	3614616	4	4	2	1
Heliopetes ericetorum	*Heliopetes laviana*	GU155986	1448470	2	3	1	1
Hemiargus ceraunus	*Leptotes pirithous*	HQ004616	1164793	2	1	3	1
Hemiargus isola	*Leptotes pirithous*	HQ004616	3066053	2	5	2	1
Hesperia	*Hesperia*	GU096951	233198	8	2	1	1

comma	*comma*						
Hesperia lindseyi	*Hesperia comma*	GU096951	373959	3	1	6	1
Hylephila phyleus	*Hylephila phyleus*	AF170859	3007334	1	6	7	2
Junonia coenia	*Junonia coenia*	GU089962	4493232	6	13	5	2
Leptotes marina	*Leptotes pirithous*	HQ004616	2266313	3	6	7	1
Lerodea eufala	*Euphyes peneia*	GU155412	1887517	2	8	1	1
Lycaeides melissa	*Lycaeides idas*	HQ004993	3492350	2	1	1	1
Lycaena cupreus	*Lycaena virgaureae*	HQ004693	882989	1	2	1	1
Lycaena editha	*Lycaena virgaureae*	HQ004693	1134456	2	1	2	1
Lycaena helloides	*Lycaena virgaureae*	HQ004693	3070416	3	5	1	1
Lycaena rubidus	*Lycaena virgaureae*	HQ004693	1756190	1	1	1	1
Lycaena xanthoides	*Lycaena virgaureae*	HQ004693	383338	2	3	3	1
Nathalis iole	*Nathalis iole*	AY954569	3746194	2	4	4	2
Nymphalis antiopa	*Nymphalis antiopa*	HQ004859	5560078	2	8	1	1
Ochlodes sylvanoides	*Ochlodes sylvanus*	HQ004883	2359390	2	1	1	1
Panoquina panoquinoides	*Panoquina lucas*	GU155448	79255	1	1	1	1
Papilio cresphontes	*Papilio cresphontes*	GU163796	2457997	1	4	1	1
Papilio eurymedon	*Papilio rutulus*	AY954560	1958385	3	3	2	1
Papilio multicaudatus	*Papilio multicaudatus*	AF044016	3130434	4	3	3	3
Papilio rutulus	*Papilio rutulus*	AY954560	2637126	5	15	8	5
Papilio zelicaon	*Papilio zelicaon*	AF044008	2629001	15	16	10	2
Phoebis agarithe	*Phoebis agarithe*	GU164486	813099	3	1	1	1
Phoebis sennae	*Phoebis sennae*	GU164658	3310792	1	1	1	1
Pholisora catullus	*Staphylus vulgata*	GU155526	4811810	1	13	4	3
Phyciodes mylitta	*Phyciodes mylitta*	AF187785	2367360	1	5	4	1
Pieris napi	*Pieris napi*	HQ004952	1858367	4	6	6	1
Pieris rapae	*Pieris rapae*	HQ004962	6072438	5	26	16	2
Plebejus acmon	*Plebejus acmon*	AF170864	539991	3	2	2	1
Plebejus saepiolus	*Plebejus saepiolus*	FJ808929	2662352	2	2	1	1
Poanes melane	*Poanes zabulon*	GU090122	376077	6	11	10	1
Polites sabuleti	*Polites themistocles*	GU090131	1709072	6	2	2	1
Pontia beckerii	*Pontia beckerii*	EU583849	2023719	2	8	5	1
Pontia occidentalis	*Pontia occidentalis*	DQ463395	2576525	2	6	5	1
Pontia protodice	*Pontia protodice*	DQ463393	5247745	2	18	12	2

Pontia sisymbrii	*Pontia beckerii*	EU583849	2529265	3	3	2	1
Pyrgus albescens	*Pyrgus communis*	GU090161	1099034	2	1	1	1
Pyrgus communis	*Pyrgus communis*	GU090161	5278007	2	9	5	1
Strymon melinus	*Strymon melinus*	GU162902	5800493	6	18	10	3
Thorybes pylades	*Thorybes pylades*	GU089293	3829092	3	4	3	1
Urbanus proteus	*Urbanus proteus*	GU155717	805402	4	3	2	1
Vanessa annabella	*Vanessa annabella*	AY788685	3022944	3	9	4	1
Vanessa atalanta	*Vanessa atalanta*	HQ005255	6277217	1	6	6	1
Vanessa cardui	*Vanessa cardui*	HQ005264	5985080	11	38	27	10
Vanessa virginiensis	*Vanessa virginiensis*	GU091474	5075311	3	1	1	1
Rhodoneura aurata	*Rhodoneura aurata*	FJ500897	N/A	N/A	N/A	N/A	N/A
Spodoptera frugiperda	*Spodoptera frugiperda*	GU090723	N/A	N/A	N/A	N/A	N/A

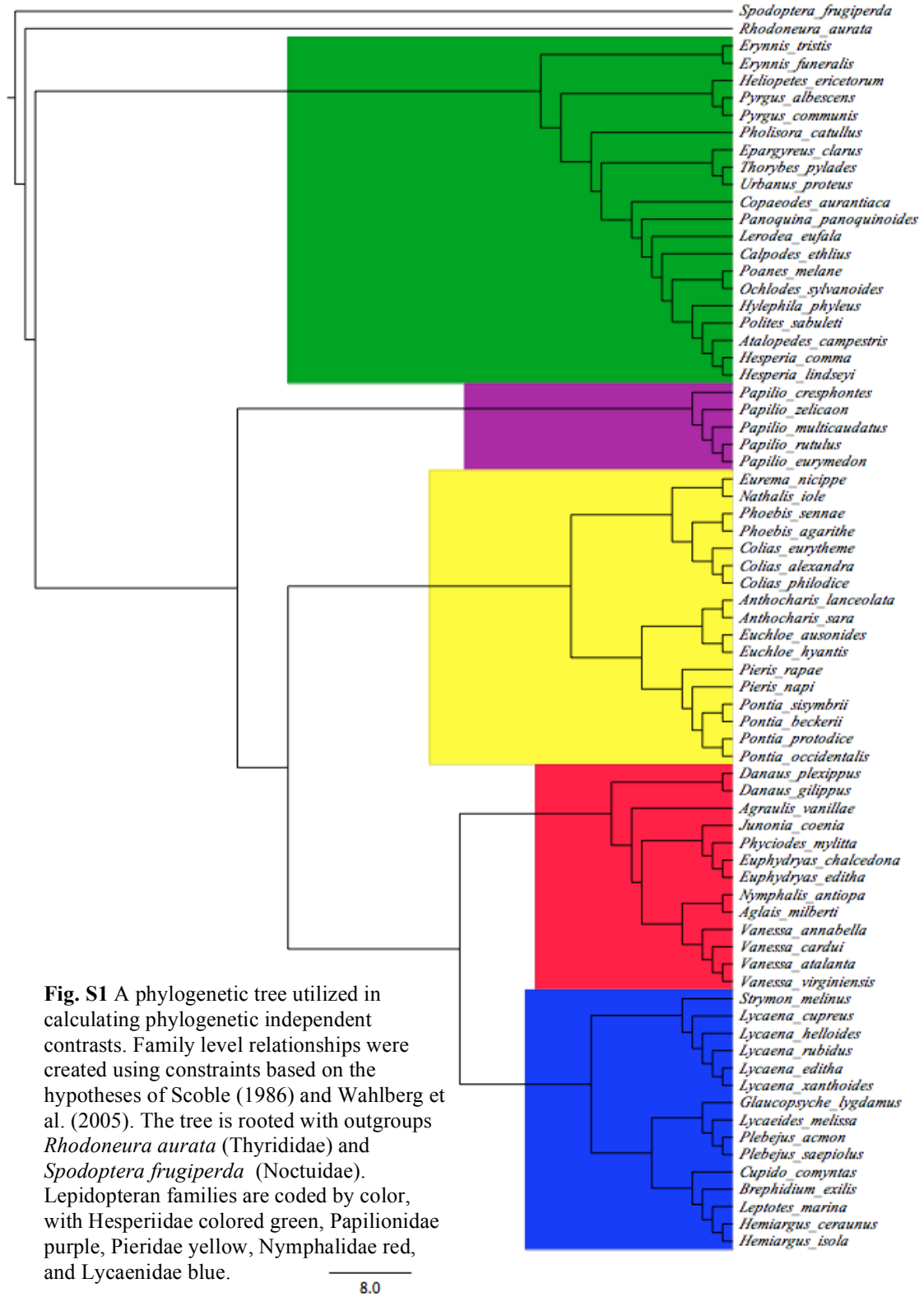


Fig. S1 A phylogenetic tree utilized in calculating phylogenetic independent contrasts. Family level relationships were created using constraints based on the hypotheses of Scoble (1986) and Wahlberg et al. (2005). The tree is rooted with outgroups *Rhodoneura aurata* (Thyrididae) and *Spodoptera frugiperda* (Noctuidae). Lepidopteran families are coded by color, with Hesperiidae colored green, Papilionidae purple, Pieridae yellow, Nymphalidae red, and Lycaenidae blue.

Appendix S2. *Criteria for selecting sequences, methods for the construction of constraint trees, GenBank accession numbers of COI sequences used, raw ecological data used for phylogenetic independent contrasts, and a figure of one phylogeny used.*

To investigate the effects of the taxonomic relatedness of exotic plants used by butterflies and skippers, four different indices were used to quantify exotic plant use. The four indices (number of species, number of genera, number of species multiplied by number of genera, and number of genera multiplied by number of families) were selected to summarize four levels of plant taxonomic relatedness. Qualitatively, all four taxonomic indices yielded similar results in linear models as response variables with geographic range and native diet breadth (log-transformed number of native genera) as predictors (Table S2).

To account for recent hypotheses of Lepidopteran phylogeny, phylogenetic independent contrasts were performed using three separate phylogenies using the number of exotic species and the number of exotic genera as separate response variables (see Appendix S1 for details on the construction of phylogenies). Qualitatively similar results were found across phylogenies and response variables, suggesting that the relationships among geographic range, native diet breadth, and exotic host use in Californian Lepidoptera hold regardless of phylogeny and method of quantifying exotic host use. For simplicity and brevity, only results from the model using the number of exotic species and the phylogeny based on the hypothesis of Scoble (1986) and Wahlberg et al. (2005) are reported in the text.

Table S2. Results from analyses using number of species, number of genera, and the following indices: species multiplied by genera (taxonomic index 1), and genera multiplied by families (taxonomic index 2) as units quantifying exotic host plant use. R^2, F, and *P* values are reported for the overall models with geographic range, native diet breadth, and Lepidopteran family (categorical variable) predicting exotic host use. Standardized beta coefficients and F values are reported for the continuous individual predictor variables (geographic range and native diet breadth) and F values are reported for the categorical variable (family). Significance of the linear model and individual variables is represented as * = $P < 0.05$; ** = $P < 0.01$; *** $P < 0.001$. The butterfly species *Vanessa cardui* was removed from the analyses using the taxonomic indices 1 and 2 as an extreme outlier.

	Model		Geographic Range		Native Diet Breadth		Family
Indice	R^2	$F_{(6,63)}$	Std. Beta	$F_{(1,63)}$	Std. Beta	$F_{(1,63)}$	$F_{(4,63)}$
Number of Species	0.36	5.96***	0.45	16.74***	0.30	8.27**	0.78
Number of Genera	0.38	6.49***	0.41	14.39***	0.38	13.22***	1.16
Taxonomic Index 1	0.34	5.26***	0.44	15.26***	0.30	7.48**	1.12
Taxonomic Index 2	0.35	5.61***	0.41	13.54***	0.32	8.95**	1.24

Table S3. Phylogenetic independent contrast results from the three phylogenies analyzed with number of exotic species and number of exotic genera used as separate response variables. The first tree was constrained based on the phylogenies of Scoble (1986) and Wahlberg et al. (2005), the second tree on the hypotheses of Mutanen et al (2010) and Regier et al. (2009), and the third tree was not constrained; see Appendix S1 for more details on constraints used. R^2, F, and *P* values are reported for the overall models with geographic range and native diet breadth predicting exotic plant species use. Standardized beta coefficients and F values are reported for the individual predictor variables (geographic range and native diet breadth). Significance of the linear model and individual variables are represented as * = $P < 0.05$; ** = $P < 0.01$; *** $P < 0.001$. For the response variable number of exotic genera using the phylogeny based on Mutanen et al. (2010) and Regier et al. (2009), the significance of the individual predictor native diet breadth was $P = 0.070$.

	Model		Geographic Range		Native Diet Breadth	
Phylogeny	R^2	$F_{(2,66)}$	Std. Beta	$F_{(1,66)}$	Std. Beta	$F_{(1,66)}$
Number of Exotic Species						
Scoble (1986); Wahlberg et al. (2005)	0.37	18.97***	0.51	26.31***	0.27	7.44**
Mutanen et al. (2010); Regier et al. (2009)	0.35	18.05***	0.51	26.10***	0.25	6.07*
No Constraint	0.44	26.44***	0.58	39.08***	0.26	7.57**
Number of Exotic Genera						
Scoble (1986); Wahlberg et al. (2005)	0.35	17.72***	0.53	28.10***	0.20	4.04*
Mutanen et al. (2010); Regier et al. (2009)	0.37	19.21***	0.56	31.64***	0.18	3.40
No Constraint	0.38	20.24***	0.56	32.15***	0.20	4.07*

LITERATURE CITED (Appendix S1 and S2)

Chew, F. S., and W. B. Watt. 2006. The green-veined white (*Pieris napi* L.), its Pierine relatives, and the systematics dilemmas of divergent character states (Lepidoptera, Pieridae). Biol. J. Linnean Soc. 88:413-435.

Mutanen, M., N. Wahlberg, and L. Kaila. 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proc. R. Soc. Lond. B. 277:2839-2848.

Pierce, N. E., M. F. Braby, A. Heath, D. J. Lohman, J. Mathew, D. B. Rand, and M. A. Travassos. 2002. The ecology and evolution of ant association in the Lycaenidae (Lepidoptera). Annu. Rev. Entomol. 47:733-771.

Pohl, N., M. P. Sison-Mangus, E. N. Yee, S. W. Liswi, and A. D. Briscoe. 2009. Impact of duplicate gene copies on phylogenetic analysis and divergence time estimates in butterflies. BMC Evol. Biol. 9:99.

Pratt, G. F., and D. M. Wright. 2002. Allozyme phylogeny of North American coppers (Lycaeninae: Lycaenidae). Pan-Pac. Entomol. 78:219-229.

Regier, J. C., A. Zwick, M. P. Cummings, A. Y. Kawahara, S. Cho, S. Weller, A. Roe, J. Baixeras, J. W. Brown, C. Parr, D. R. Davis, M. Epstein, W. Hallwachs, A. Hausmann, D. H. Janzen, I. J. Kitching, M. A. Solis, S.-H. Yen, A. L. Bazinet, and C. Mitter. 2009. Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. BMC Evol. Biol. 9:280.

Scoble, M. J. 1986. The structure and affinities of the Hedyloidea: a new concept of the butterflies. Bull. Br. Mus. Nat. Hist., Entomol. Ser. 53:251-286.

Wahlberg, N., M. F. Braby, A. V. Z. Brower, R. de Jong, M.-M. Lee, S. Nylin, N. E. Pierce, F. A. H. Sperling, R. Vila, A. D. Warren, and E. Zakharov. 2005a. Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. Proc. R. Soc. Lond. B. 272:1577-1586.

Wahlberg, N., A. V. Z. Brower, and S. Nylin. 2005b. Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae). Biol. J. Linnean Soc. 86:227-251.

Wahlberg, N., E. Weingartner, and S. Nylin. 2003. Towards a better understanding of the higher systematics of the Nymphalidae (Lepidoptera: Papilionoidea). Mol. Phylogenet. Evol. 28:473-484.

Warren, A. D., J. R. Ogawa, and A. V. Z. Brower. 2008. Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperiidae (Lepidoptera: Hesperioidea). Cladistics 24:1-35.

Zakharov, E. V., M. S. Caterino, F. A. H. Sperling. 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). Syst. Biol. 53:193-215.