The global distribution of diet breadth in insect herbivores

Matthew L. Forister[a,1], Vojtech Novotny[b,c,d], Anna K. Panorska[e], Leontine Baje[d,f], Yves Basset[b,g], Philip T. Butterill[b,c], Lukas Cizek[b,c], Phyllis D. Coley[g,h], Francesca Dem[d,i], Ivone R. Diniz[j], Pavel Drozd[k], Mark Fox[l], Andrea E. Glassmire[a], Rebecca Hazen[l], Jan Hrcek[b,c,m], Joshua P. Jahner[a], Ondrej Kaman[b,c], Tomasz J. Kozubowski[e], Thomas A. Kursar[g,h], Owen T. Lewis[m], John Lill[n], Robert J. Marquis[o], Scott E. Miller[p], Helena C. Morais[j], Masashi Murakami[q], Herbert Nickel[r], Nicholas A. Pardikes[a], Robert E. Ricklefs[o,1], Michael S. Singer[s], Angela M. Smilanich[a], John O. Stireman[t,u], Santiago Villamarín-Cortez[u], Stepan Vodka[b,c], Martin Volf[b,c], David L. Wagner[v], Thomas Walla[u,w], George D. Weiblen[x], and Lee A. Dyer[a,u]

[a]Department of Biology, Program in Ecology, Evolution and Conservation Biology and [e]Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557; [b]Faculty of Science, University of South Bohemia, 370 05 Ceske Budejovice, Czech Republic; [c]Biology Centre, Czech Academy of Science, 370 05 Ceske Budejovice, Czech Republic; [d]New Guinea Binatang Research Center, Madang, Papua New Guinea; [f]Department of Biology, University of Papua New Guinea, Port Moresby, Papua New Guinea; [g]Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancon, Panama; [h]Department of Biology, University of Utah, Salt Lake City, UT 84112-0840; [i]Environmental Futures Research Institute, Griffith University, Queensland 4111, Australia; [j]Departamento de Zoologia, Universidade de Brasília, Brasília, Distrito Federal, 70 910-900; [k]Department of Biology, University of Ostrava, 710 00 Ostrava, Czech Republic; [l]Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, LA 70118; [m]Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom; [n]Department of Biological Sciences, George Washington University, Washington, DC 20052; [o]Department of Biology and Whitney R. Harris World Ecology Center, University of Missouri–St. Louis, St. Louis, MO 63121; [p]National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012; [q]Department of Biology, Chiba University, Chiba 263-8522, Japan; [r]J. F. Blumenbach Institute of Zoology and Anthropology, Göttingen University, 37073 Göttingen, Germany; [s]Department of Biology, Wesleyan University, Middletown, CT 06459; [t]Department of Biological Sciences, Wright State University, Dayton, OH 45435; [u]Sección Invertebrados, Museo Ecuatoriano de Ciencias Naturales, Quito, Ecuador; [v]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269; [w]Department of Biology, Colorado Mesa University, Grand Junction, CO 81507; and [x]Bell Museum and Department of Plant Biology, University of Minnesota, St. Paul, MN 55108-1095

Understanding variation in resource specialization is important for progress on issues that include coevolution, community assembly, ecosystem processes, and the latitudinal gradient of species richness. Herbivorous insects are useful models for studying resource specialization, and the interaction between plants and herbivorous insects is one of the most common and consequential ecological associations on the planet. However, uncertainty persists regarding fundamental features of herbivore diet breadth, including its relationship to latitude and plant species richness. Here, we use a global dataset to investigate host range for over 7,500 insect herbivore species covering a wide taxonomic breadth and interacting with more than 2,000 species of plants in 165 families. We ask whether relatively specialized and generalized herbivores represent a dichotomy rather than a continuum from few to many host families and species attacked and whether diet breadth changes with increasing plant species richness toward the tropics. Across geographic regions and taxonomic subsets of the data, we find that the distribution of diet breadth is fit well by a discrete, truncated Pareto power law characterized by the predominance of specialized herbivores and a long, thin tail of more generalized species. Both the taxonomic and phylogenetic distributions of diet breadth shift globally with latitude, consistent with a higher frequency of specialized insects in tropical regions. We also find that more diverse lineages of plants support assemblages of relatively more specialized herbivores and that the global distribution of plant diversity contributes to but does not fully explain the latitudinal gradient in insect herbivore specialization.

host range | latitudinal gradient | niche width | Pareto distribution | specialization

Variation in dietary specialization among individuals, populations, and species drives numerous ecological and evolutionary processes. Differences in diet breadth and composition mediate the coexistence of competitors (1), the persistence of species in the face of environmental disturbance (2), the diversity of interactions, and the stability of entire networks of interacting species (3, 4). At the ecosystem level, the top-down effect of predators on primary productivity can be controlled by the level of herbivore specialization (5). At the scale of evolutionary

diversification, differences among lineages in rates of speciation and extinction can be understood in terms of variation in dietary specialization (6). Progress in addressing all of these issues has been limited by disparity in the methods used to quantify specialization (7) and the rarity of comparative datasets encompassing large numbers of species and regions (8).

Herbivorous insects have long served as models for the study of specialization (9), because they use a variety of plant resources

Significance

Dietary specialization determines an organism's resource base as well as impacts on host or prey species. There are important basic and applied reasons to ask why some animals have narrow diets and others are more generalized, and if different regions of the Earth support more specialized interactions. We investigated site-specific host records for more than 7,500 species of insect herbivores. Although host specialists predominate, the proportion of specialists is affected by the diversity of hosts and shifts globally, supporting predictions of more exclusive tropical interactions. These results not only affect our understanding of the ecology of food webs, but also have implications for how they respond to environmental change, as well as for ecosystem management and restoration.

in many different ways, and their host plants are discrete resources; thus, diet breadth of a given herbivore species can be quantified as the number of plant taxa that it eats. In addition to serving as models for the study of diet breadth, herbivorous insects are important in their own right as one of the most abundant and diverse forms of life on Earth, and the consumption of plant material by insect herbivores is a dominant mover of energy and matter through terrestrial ecosystems (10). Although herbivorous insects are known collectively for narrow diet breadth, species vary significantly in the number of host plant taxa that they attack, and this variation has scarcely been described in a quantitative manner (11). It is unclear, for example, whether species with relatively specialized and generalized diets anchor the ends of a continuum or if diet breadth is discontinuous, with specialists and generalists forming distinct modes. Researchers have most commonly treated diet breadth as bimodal, distinguishing only generalists and specialists (7). In some cases, this distinction is a simplification for theoretical discussion (12), and in other cases it is simply a convenience (13, 14).

A quantitative description of the relative frequencies of specialists and generalists is important for understanding the evolution and ecology of plant–insect interactions. Because specialists can evolve from generalists and vice versa (6), species presumably pass through stages of intermediate diet breadth; thus, a bimodal distribution of the number of hosts attacked would suggest that intermediate levels of diet breadth are maladaptive. Similarly, if the distribution of diet breadth is continuous, the relative frequencies of herbivores in different diet breadth categories could be examined in the context of environmental determinants of niche width. It has also been suggested that treating herbivores as either specialists or generalists creates errors in estimates of the diversity of ecological communities, including the number of species of arthropods on Earth (15). Here, we quantify herbivore host ranges from field collections of larval and adult insects. We focus many of our analyses on larval Lepidoptera (caterpillars), for which we have the most geographically extensive data, but also include herbivorous insects from other feeding guilds as well as limited samples of parasitoids (flies and wasps) that feed on caterpillars. Parasitoids potentially offer an informative contrast with insect herbivores, because they are also highly specialized but feed at a higher trophic level (16).

Our global dataset of plant–insect interactions encompasses thousands of species of herbivores from 17 localities spanning 63° latitude from Canada to Brazil in the Western Hemisphere and from the United Kingdom and Japan to Papua New Guinea in the Eastern Hemisphere (Fig. 1*A* and *SI Appendix*, section S1). In addition to our primary objective of characterizing the distribution of specialists and generalists among insect herbivores, the latitudinal breadth of our samples allows us to address hypotheses concerning global patterns of ecological interactions. In particular, several explanations for the higher diversity of species at lower latitudes assume that interactions in tropical regions are more specialized (17).

For example, a greater presence of specialists at lower latitudes could promote diversification of plant traits (18), and herbivory by specialists can contribute to coexistence among competing plants (19). However, evidence both for and against global gradients in interaction specificity has been reported (20–22), and we do not know if latitudinal trends in specificity might result from heterogeneity of resources or other factors.

Results

The distribution of taxonomic diet breadth is characterized by a highly skewed, concave distribution that is well-fit by a discrete, truncated Pareto power law. Examples of diet breadth distributions for both Lepidoptera and other herbivores are shown in Fig. 1 *B* and *C*. Also shown in Fig. 1 are visualizations of fit to other distributions that are commonly used for ecological count
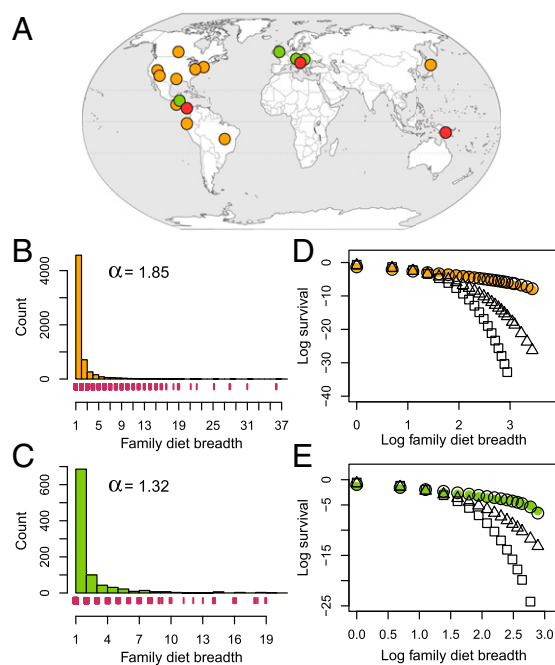

Fig. 1. (*A*) Study locales and the distribution of diet breadth for (*B*) Lepidoptera and (*C*) all other herbivores. Points on the globe are shown in orange for Lepidoptera study sites, green for other herbivore study sites, and red for study sites for both. Histograms in *B* and *C* illustrate counts of the numbers of herbivores associated with different numbers of host plant families; also shown is the shape parameter (α) from the discrete, truncated Pareto distribution. Tick marks under histograms indicate individual observations for visualization in the thin tail of the distributions. (*D* and *E*) Survival plots illustrate the fit of the Pareto (white circles), geometric (triangles), and Poisson (squares) distributions to the data (colored symbols). Log survival on the y axis is $\ln(P(X > x))$, which is the natural logarithm of the probability of herbivores having a greater diet breadth (*X*) than the corresponding value (*x*) on the x axis. Note that most analyses focus on diet breadth at the scale of individual sites; for simplicity, diet breadth is shown here across sites.

data (Fig. 1 *D* and *E*). Compared with the truncated Pareto distribution, the geometric and Poisson distributions fit the diet breadth data poorly. We use the discrete, truncated Pareto, because the long tail of the empirical distributions of diet breadth (Fig. 1*B*) suggests a power law, and the discrete, truncated formulation is appropriate to the particular data being modeled (the number of hosts attacked by herbivores is a discrete count, and truncation results from the number of plant taxa attacked by any one herbivore always being less than the total number of plant taxa sampled). Furthermore, the shape parameter (α) of the discrete, truncated Pareto is more informative than measures of central tendency, such as the mean, for highly nonsymmetrical distributions. Higher values of α indicate a greater proportion of more specialized herbivores.

The distributions in Fig. 1 *B* and *C* illustrate family-level diet breadth (the number of host plant families attacked), which is highly correlated with species-level diet breadth (the number of host plant species attacked): Pearson's correlation coefficient between family- and species-level diet breadth across all herbivores = 0.89 ($P < 0.001$) (*SI Appendix*, section S2). Species-level diet breadth is also closely fit by the discrete, truncated Pareto (*SI Appendix*, Fig. S2); high values of α, indicating a concentration of specialists and a long tail of generalists, characterize most of the taxonomic and functional subsets of the global data (*SI Appendix*, Table S3). Based on these data, an average of 69% of caterpillar species are associated with a single host family at each site; above 25° latitude, this value is 60%, and at sites less than

25° latitude, it is 83%. Among the sampled guilds of herbivorous insects, 76% of species associate with a single host family, although values for individual guilds vary widely (species-level diet breadth for the different feeding guilds as well as Lepidoptera shows similar patterns) (*SI Appendix*, Table S3).

Geographically, the discrete, truncated Pareto shows a consistently good fit across latitudes (*SI Appendix*, Fig. S4), whereas the shape parameter (α) increases toward the equator ($P = 0.0046$; $F_{2,10} = 8.87$; $R^2 = 0.64$) (Fig. 2A). The increase in α toward the equator for larval Lepidoptera corresponds to a greater relative frequency of specialized herbivores. In contrast, maximum diet breadth does not change with latitude, although a greater number of potential hosts is available in the tropics (white symbols in Fig. 2B). Although the upper limit of the distribution is unchanged with latitude, there is a shift throughout the distribution toward more specialized diets, such that most insect herbivores become more specialized toward the equator, which is illustrated in Fig. 2B (*SI Appendix*, Figs. S5 and S6).

Taxonomic diet breadth is a convenient but incomplete index of host range or dietary niche width, because herbivores attack particular plants for a variety of reasons (e.g., phytochemistry or geography) that may or may not be captured by simply counting the number of species eaten. As a complementary approach to investigating the latitudinal gradient in specialization, we calculated phylogenetic diet breadth (23) for Lepidoptera species associated with angiosperms, for which a robust phylogeny is available at the family level (24). Phylogenetic family-level diet breadth, as measured by phylogenetic distance (PD) among hosts, changes globally: mean PD declines toward lower latitudes ($P < 0.001$; $F_{2,10} = 11.58$; $R^2 = 0.70$) (Fig. 2C), despite the greater phylogenetic diversity of tropical compared with temperate plants (25). This result is driven by the increase in the number of specialists (species with low PD) at lower latitudes. Thus, when specialists (i.e., herbivores associated with a single host family) are removed from the analysis, PD does not change with latitude (white symbols in Fig. 2C).

Previously, insect herbivores were reported to be more specialized at lower latitudes in the Western Hemisphere (21), whereas no latitudinal trend in specialization was found in the Eastern Hemisphere, where the sampled vegetation had been standardized for latitudinal differences in phylogenetic diversity (20). Results from this analysis for both taxonomic and phylogenetic diet breadth suggest that data from both hemispheres conform to a single global latitudinal gradient (Fig. 2 A and C). These analyses differ from previous studies in a number of ways: we have focused on a single life history stage (larvae), we have included phylogenetic information for insect hosts from all sites, and we have increased the number of sites from both hemispheres. The inference of a global gradient in herbivore specialization is supported by temperate and tropical comparisons among different feeding guilds, which is illustrated by variation in α for family- and species-level diet breadth in Fig. 3. Average α across samples of species-level diet breadth rarefied to the lowest number of hosts in each comparison is significantly higher in tropical than temperate samples (Wilcoxon signed rank test, $P = 0.016$), and the same is true for family-level diet breadth if the one apparent exception, phloem suckers, is removed ($P = 0.031$; without phloem suckers removed, $P = 0.16$). Guilds with the most intimate host plant associations (gallers and leaf miners) show the highest levels of specialization (26), and it is noteworthy that such interguild differences are evident, despite latitudinal variation (Fig. 3). Among the most generalized guilds are mobile adult chewers, such as leaf beetles (family Chrysomelidae), that are able to move among individual plants to a greater extent than sessile or wingless life stages and may achieve a broader diet than larvae through selection of plant tissues that limit exposure to phytochemicals (27).

Variation in specialization among lineages and regions is addressed by theories of adaptive radiation and coevolution (28), which predict a negative relationship between the diversity of available resources and the diet breadth of consumers. Unexploited resources, for example, allowed the Hawaiian honeycreepers to evolve into a large number of specialized species (29). Thus, for herbivorous insects, it has been hypothesized that lower latitudes might harbor more specialized herbivores because of the diversity of hosts; similarly, the most species-rich lineages of plants might allow for the evolution of a greater number of specialists (30–32). However, these patterns have been difficult to evaluate empirically given the scope of many previous studies. Based on more than 6,000 species of Lepidoptera, we investigated connections between herbivore diet breadth and plant diversity and found that host plant families with high global species richness are associated with assemblages of herbivores with relatively narrow species-level diet breadth ($P < 0.001$; $R^2 = 0.56$) (Fig. 2D). This relationship holds when linear models contain covariates associated with plant families, including sample size (the number of experimental rearings from a family), phylogenetic age, and latitudinal extent of geographic range (*SI Appendix*, Table S4 shows analyses of latitudinal subsets of herbivores). The relationship between host richness and dietary specialization could inform our understanding of the latitudinal gradient in herbivore specialization if lower latitudes contain more diverse lineages of hosts. This possibility was addressed with path analysis, including a direct effect of latitude
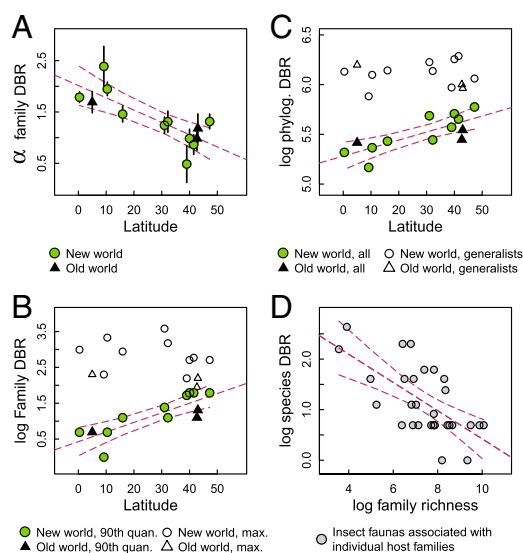

Fig. 2. Patterns in the distribution of diet breadth (DBR) for Lepidoptera. (A) Latitudinal trend in the shape parameter (α) for family-level DBR with bootstrapped SEs for individual sites; larger values of α indicate distributions with a higher density of herbivores having narrow DBRs. (B) Maximum observed DBR (white symbols) and DBR in the 90th quantile of the DBR distributions (colored symbols) vs. latitude. Lower values of the 90th quantile correspond to a distribution in which more herbivores are compressed toward lower, more specialized DBRs. (C) Latitudinal patterns in phylogenetic DBR among sites for all species are shown with colored symbols, and more generalized species that attack more than one host family are shown with white symbols (log units on the y axis are in millions of years). (D) Relationship between median species-level DBR (as the number of host species attacked; y axis) for herbivores associated with particular families and the species richness of the plant families (x axis); more species-rich plant families host more specialized herbivores (*SI Appendix*, section S4 shows analyses on latitudinal subsets of herbivores). For all panels, 95% confidence limits are shown around linear relationships, and the natural logarithm is used in all cases. In A–C, circles (white and green) are New World sites, and triangles (white and black) are Old World sites; differences between white and colored symbols in B and C are specific to B and C.
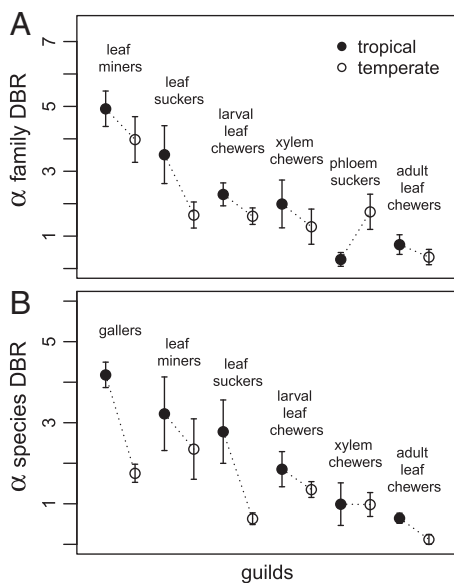
Fig. 3. Diet breadth (DBR) comparisons for herbivore guilds from tropical and temperate communities for (*A*) family-level DBR and (*B*) species-level DBR. Higher values of the shape parameter (α) indicate more specialized diets. Means and SDs are based on rarefaction to the lowest number of host taxa sampled in each pairwise comparison connected by dotted lines. More than one community is represented by some but not all of the points (*SI Appendix*, section S1). Note that not all guilds could be analyzed for both species- and family-level DBR (*Materials and Methods* and *SI Appendix* contain more details).

on specialization as well as an indirect effect of latitude mediated through plant richness (the number of plant families and species encompassed by insect sampling at each site). We found that the effect of latitudinal variation in plant richness on specialization is approximately one-fourth the direct effect of latitude on the global trend in specialization (*SI Appendix*, Fig. S7), indicating that plant diversity may contribute to variation in specialization but does not fully explain the global pattern of interactions (33).

Discussion

In summary, the distribution of diet breadth for insect herbivores conforms to a power law, with a majority of species associated locally with a single plant taxonomic family or species and a long tail of more generalized herbivores. The distribution of specialization shifts globally with latitude (Fig. 2*A*), which confirms the long-standing expectation that interactions are more specific at lower latitudes (17) and contributes to our understanding of the latitudinal diversity gradient. Plant diversity affects insect diet breadth (Fig. 2*D* and *SI Appendix*, Fig. S7), which may, in turn, feed back onto plant diversity through either coevolutionary (30) or ecological (12, 34) interactions, although the investigation of such feedback processes is beyond the scope of this study (35).

When interpreting the frequencies of insects in different diet breadth categories (Fig. 1), it is important to note that we have focused on local estimates of diet breadth derived from collections and rearings at focal sites. The alternative approach (using host records that encompass entire geographic ranges of herbivores) has less relevance for understanding ecological determinants of diet breadth (36) and faces the additional challenge that widespread generalists may often be composed of cryptic, localized specialists (37). Clearly, the distribution of herbivore diet breadth is continuous (specialists and generalists are not separated by a gap in frequency distributions) (Fig. 1), although it may be useful to consider herbivores as belonging to either the great majority of specialists or the long, thin tail of generalists. A

similar frequency distribution is apparent in the diet breadth of insect parasitoids (*SI Appendix*, Fig. S3), suggesting that the highly concave, skewed distribution could be inherent to the parasitic habit, of which insect herbivores are only one example, albeit the most well-studied (38). In general, the predominance of specialists is relevant to issues in natural resource management and challenges the idea that, in human-impacted systems, the interactions among novel suites of co-occurring species will be comparable in structure with less-degraded systems (39, 40).

Recent approaches to studying biotic networks include comparisons of distributions of natural interactions with theoretical and mathematical predictions. In randomly assembled networks of interactions, the frequency of highly connected nodes drops rapidly beyond the mode, such as in a Poisson distribution (41, 42), and some natural networks (for example, plant–pollinator interactions) deviate from that pattern by having a small portion of relatively overconnected nodes, which we have observed with generalist herbivores (41, 43, 44). For these more heterogeneous networks, it has been suggested that facilitating processes underlie the long, thin tail of the frequency distribution (45). In the case of herbivore diet breadth, the process of host range expansion could be such a process: initial expansion of diet breadth might be rare, but after more than one host plant family has been colonized, adaptations for generalized feeding could facilitate the colonization of new hosts (46). This possibility is consistent with the observation that highly polyphagous species of butterflies are more likely to use novel hosts (47, 48), and the ubiquity of the highly peaked, skewed distribution of diet breadth suggests that any processes facilitating dietary generalization operate not only in different regions of the globe, but also within different lineages and guilds of herbivores. An additional mechanism for the observed frequencies of diet breadth could be disruptive selection, favoring either extreme specialization or increasing generalization. However, the biological factors that favor dietary generalism as an ever-present but relatively infrequent life history strategy in herbivorous insects must await additional study.

Materials and Methods

Data Collection. Rearing of field-collected caterpillars (Lepidoptera) to establish consumer–host relationships were undertaken at 13 sites in North, Central, and South America, Papua New Guinea, Japan, and central Europe (Fig. 1*A* and *SI Appendix*, Table S1). Collections and rearings focused on externally feeding caterpillars, mostly macrolepidopterans, with broader sampling at some sites. Site-specific details are provided in *SI Appendix* for each dataset and site or a collection of sites when more than one site has been managed in a similar way. We tabulated data for other herbivores to study variation in diet breadth within and among ecological guilds, with the goal of encompassing both sucking and chewing feeding modes, species using various plant resources (including leaves, phloem sap, and wood), and species having diverse life histories, from species intimately connected with plants, such as miners and gallers, to those that are mobile, either as larvae or adults (Fig. 3). Data were compiled from seven sites (three sites in common with the Lepidoptera data and four additional sites) (Fig. 1*A* and *SI Appendix*, Table S2). As with Lepidoptera, all data for these herbivores are based on field collections for both larval and adult stages (depending on the herbivore taxa).

Heterogeneity among sites in methodology and sampling effort is almost always present in global datasets, which emphasizes the importance of statistical controls (covariates) to account for variation in sampling effort. Although this approach has a well-established history in ecology, covariates introduce complexity to models and reduce degrees of freedom for main effects; however, substantial power is gained by the inclusion of additional sites, despite differences in design. Variation in sampling effort could be particularly important when considering variation in diet breadth: an understudied herbivore community might seem more specialized, because not all plant–insect interactions will have been sampled, or less specialized if rare species are mostly specialists (and missed by sampling). Furthermore, the discovery of interactions will necessarily take more time in areas with more diverse floras and faunas. However, our most diverse sites are tropical and extremely well-sampled, with 71 site-years of sampling below 30° latitude just for Lepidoptera (*SI Appendix*, Table S1). We use the number of

years of study along with other appropriate covariates (including the number of rearing records and the geographical area of study sites) to address sampling issues throughout our analyses, which are described below and in *SI Appendix*. Ultimately, biological signal is evident at multiple scales (in latitudinal patterns and the consistent fit of the Pareto distribution among sites), despite differences in methodologies and sampling effort. It can also be noted that the number of rearing records per Lepidoptera species does not vary with latitude ($F_{1,11} = 0.84$; $R^2 = 0.0071$; $P = 0.38$). Although we have been successful in detecting relationships among our focal variables, future research in the area of global interactions could reasonably strive for a greater number of sites with standardized methodologies that would potentially increase explained variance in analyses, such as the path models in *SI Appendix*, Fig. S7.

To determine host associations for each herbivore species, data from each site were subjected to site-specific preanalysis filters appropriate to the methods of collection and taxonomic knowledge at each site (*SI Appendix*, section S1). Data for Lepidoptera from some of the sites were analyzed previously in the works by Dyer et al. (21) and Novotny et al. (20). Here, we updated those previous datasets with additional years (more than double the previously examined number of years for some sites) and added new locations for Lepidoptera (Ohio, Great Basin, and Japan). We also added 18 datasets for other herbivores that had not previously been brought together into one analysis (*SI Appendix*, section S1). Identification procedures varied among sites, but voucher specimens have been deposited at appropriate museums, and both insect and plant species were identified by knowledgeable taxonomists in so far as possible; several sites additionally used DNA barcoding to validate species identifications.

Pareto Distribution. Because the distribution of diet breadth is highly skewed (Fig. 1), the core of our statistical approach is the use of a theoretical distribution appropriate to the structure of the data. We have used a version of the Pareto distribution, from which a shape parameter (α) can be extracted that serves as a useful summary statistic; higher values of α correspond to diet breadth distributions that are more highly peaked, with a greater density of specialists. The Pareto distribution (49) is widely used in a number of fields of science, and the truncated Pareto has been proposed as an important extension (50). Here, we use a form of the Pareto that is truncated and discrete, and thus, it is appropriate for ecological count data (51); in our case, we are interested in counts of host plant families and host plant species associated with specific herbivore species. We say that a random variable X has a truncated discrete Pareto distribution with parameters α, β, and γ if

$$P(X = x) = \frac{\frac{1}{x^\alpha} - \frac{1}{(x+1)^\alpha}}{\frac{1}{\gamma^\alpha} - \frac{1}{(\beta+1)^\alpha}}$$

for any natural number of host plant taxa (x), such that $\gamma \leq x \leq \beta$. Here, α is any real number, and γ and β are positive integers. Parameters γ and β are the lower and upper truncation parameters, respectively, for which maximum likelihood estimates were used: the sample minimum for γ and the sample maximum for β. We estimated α by minimizing the sum of squares of the differences between the model survival function and the empirical survival function. The distribution was fit to data using custom R scripts (*SI Appendix*, section S5).

As a measure of goodness of fit or closeness of the modeled discrete truncated Pareto distribution to the empirical distribution of the data, we used the maximum absolute difference between the modeled and the empirical cumulative distribution functions. This approach provides the largest difference between probabilities computed using the model and the relative frequency from the data, and it gives an intuitive notion of closeness (52). Fits of the discrete, truncated Pareto to our data were visualized using plots of survival functions, as in the work by Aban et al. (50), which are useful for comparing values in the tail of the distribution with predictions generated by theoretical distributions (53–55). We compared the discrete, truncated Pareto with the geometric and Poisson distributions; in Fig. 1, it can be seen that the Pareto (white circles) predict the observations (colored symbols) throughout the tail of the distribution.

Relationships between parameters from the Pareto distribution fitted to diet breadth (e.g., the shape parameter, α) and latitude were investigated for Lepidoptera with linear models that included sample size (the number of experimental rearings per site), area (of sampled locations), the number of herbivore species, and the number of years (over which sampling was conducted) as covariates. For other herbivores, we did not have the same well-sampled latitudinal gradient but were able to make pairwise comparisons between temperate and tropical sites (*SI Appendix*, Table S2). Because

these analyses did not involve multiple regression models, we did not include covariates associated with sampling effort or other factors as in the Lepidoptera analyses. Instead, we used a rarefaction approach based on the number of plant taxa (families and species) sampled, which is a key axis along which datasets differ: when one dataset includes 15 hosts and another includes only 10 hosts, the latter will potentially be biased toward a more narrow observed diet breadth among herbivores. For each set of comparisons (for example, matrices of plant–insect associations for leaf miners from tropical and temperate communities), we subsampled each matrix down to the lowest number of plant taxa sampled for any of the matrices involved in the comparison, and we did this 1,000 times using the sample function in R and inspected means and variances across resampled replicates.

Phylogenetic Analyses. To investigate phylogenetic diet breadth of Lepidoptera species, we used PD among host plant families calculated with the R package picante and using the angiosperm phylogeny from the work by Davies et al. (24), which was the most complete angiosperm phylogeny available at the time of analysis (nonangiosperm host records were excluded from these analyses). As with parameters from the fitted Pareto distributions (see above), the relationship between PD and latitude was investigated using linear models and covariates to account for sampling effort.

Plant Diversity. Diversity of resources (host plants) can affect the evolution of consumer traits by providing opportunities for local adaptation and specialization, a possibility that we have investigated in two complementary ways: using plant lineages (for which resource diversity is global species richness of plant families) followed by using geographic sites as replicates (for which resource diversity is richness of sampled plant taxa). For the first approach, multiple regression used median species-level diet breadth of insect herbivores as the response variable and the following independent variables: species richness (number of species within plant families), relative age (extracted from the angiosperm phylogeny) (24), latitudinal range, and sample size (the number of times in which an insect was reared from any species in a host plant family). Median diet breadth was used here instead of α from the Pareto, because comparisons were being made among subsets of data (insects associated with plant families) that varied greatly in sample size, and the Pareto could not always be fit to the smallest sample sizes. However, the use of the median is conservative, because shifts in the tail of the distribution may not be reflected in the median value when comparing two distributions that have similar numbers of extreme specialists. Species richness (the total number of species within plant families) for these analyses was taken from the angiosperm diversity website maintained by the Missouri Botanical Garden (www.mobot.org/MOBOT/research/APweb/), which was also used as a reference to standardize family names from the different databases used in this project. Latitudinal range (the maximum extent of north–south latitude occupied by any species in the plant family) was taken from family-level distribution maps in Heywood (56). Only plant families for which all measures were available (richness, relative age, and latitudinal range) were included in analyses. Furthermore, a sample size cutoff was used to avoid plant families that were less well-characterized from the perspective of insect rearings: we considered families from which insects had been reared at least 100 times. To investigate the robustness of results, the multiple regression using these variables was repeated with subsets of the data, specifically only using insects for which rearing records were restricted to sites either greater than 25° or less than or equal to 25° latitude.

Our second approach to understanding the influence of plant diversity on dietary specialization involved path analyses and variation among sites in dietary specialization and plant richness. Path analysis is useful in this context, because it allows for the simultaneous analysis of direct and indirect effects. Specifically, path analyses included the α shape parameter from the Pareto distributions and plant richness per site as endogenous variables (plant richness for each site is the number of plant families and species associated with the sampled herbivores). The exogenous variable was simply latitude, which pointed directly to α and indirectly to α through plant richness (allowing for the possibility that latitude affects α through resource diversity but also has effects that are not explained by global variation in plant richness). The model was evaluated using plant richness as both the number of plant species and the number of plant families sampled at sites.

1. Büchi L, Vuilleumier S (2014) Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. *Am Nat* 183(5):612–624.
2. Devictor V, Julliard R, Jiguet F (2008) Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos* 117(4): 507–514.
3. Mougi A, Kondoh M (2012) Diversity of interaction types and ecological community stability. *Science* 337(6092):349–351.
4. Lever JJ, van Nes EH, Scheffer M, Bascompte J (2014) The sudden collapse of pollinator communities. *Ecol Lett* 17(3):350–359.
5. Singer MS, et al. (2014) Herbivore diet breadth mediates the cascading effects of carnivores in food webs. *Proc Natl Acad Sci USA* 111(26):9521–9526.
6. Hardy NB, Otto SP (2014) Specialization and generalization in the diversification of phytophagous insects: Tests of the musical chairs and oscillation hypotheses. *Proc Biol Sci* 281(1795):20132960.
7. Devictor V, et al. (2010) Defining and measuring ecological specialization. *J Appl Ecol* 47(1):15–25.
8. Forister ML, Dyer LA, Singer MS, Stireman JO, 3rd, Lill JT (2012) Revisiting the evolution of ecological specialization, with emphasis on insect-plant interactions. *Ecology* 93(5):981–991.
9. Futuyma DJ, Moreno G (1988) The evolution of ecological specialization. *Annu Rev Ecol Syst* 19:207–233.
10. Seastedt T, Crossley D, Jr (1984) The influence of arthropods on ecosystems. *Bioscience* 34(3):157–161.
11. Jaenike J (1990) Host specialization in phytophagous insects. *Annu Rev Ecol Syst* 21: 243–273.
12. Janzen DH (1970) Herbivores and the number of tree species in tropical forests. *Am Nat* 104(940):501–528.
13. Erwin TL (1982) Tropical forests: Their richness in Coleoptera and other arthropod species. *Coleopt Bull* 36(1):74–75.
14. Dyer LA (1995) Tasty generalists and nasty specialists? Antipredator mechanisms in tropical lepidopteran larvae. *Ecology* 76(5):1483–1496.
15. Ødegaard F (2000) How many species of arthropods? Erwin's estimate revised. *Biol J Linn Soc Lond* 71(4):583–597.
16. Wiegmann BM, Mitter C, Farrell B (1993) Diversification of carnivorous parasitic insects: Extraordinary radiation or specialized dead end? *Am Nat* 134(5):737–754.
17. Schemske DW, Mittelbach GG, Cornell HV, Sobel JM, Roy K (2009) Is there a latitudinal gradient in the importance of biotic interactions? *Annu Rev Ecol Evol Syst* 40:245–269.
18. Agrawal AA, Fishbein M (2008) Phylogenetic escalation and decline of plant defense strategies. *Proc Natl Acad Sci USA* 105(29):10057–10060.
19. Dyer LA, Letourneau DK, Chavarria GV, Amoretti DS (2010) Herbivores on a dominant understory shrub increase local plant diversity in rain forest communities. *Ecology* 91(12):3707–3718.
20. Novotny V, et al. (2006) Why are there so many species of herbivorous insects in tropical rainforests? *Science* 313(5790):1115–1118.
21. Dyer LA, et al. (2007) Host specificity of Lepidoptera in tropical and temperate forests. *Nature* 448(7154):696–699.
22. Morris RJ, Gripenberg S, Lewis OT, Roslin T (2014) Antagonistic interaction networks are structured independently of latitude and host guild. *Ecol Lett* 17(3):340–349.
23. Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annu Rev Ecol Syst* 33:475–505.
24. Davies TJ, et al. (2004) Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proc Natl Acad Sci USA* 101(7):1904–1909.
25. Kerkhoff AJ, Moriarty PE, Weiser MD (2014) The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc Natl Acad Sci USA* 111(22):8125–8130.
26. Novotny V, et al. (2010) Guild-specific patterns of species richness and host specialization in plant-herbivore food webs from a tropical forest. *J Anim Ecol* 79(6): 1193–1203.
27. Wäckers FL, Romeis J, van Rijn P (2007) Nectar and pollen feeding by insect herbivores and implications for multitrophic interactions. *Annu Rev Entomol* 52:301–323.
28. Ackermann M, Doebeli M (2004) Evolution of niche width and adaptive diversification. *Evolution* 58(12):2599–2612.
29. Lovette IJ, Bermingham E, Ricklefs RE (2002) Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proc Biol Sci* 269(1486): 37–42.
30. Ehrlich PR, Raven PH (1964) Butterflies and plants: A study in coevolution. *Evolution* 18(4):586–608.
31. Janz N, Nylin S, Wahlberg N (2006) Diversity begets diversity: Host expansions and the diversification of plant-feeding insects. *BMC Evol Biol* 6(1):4.
32. Fordyce JA (2010) Host shifts and evolutionary radiations of butterflies. *Proc Biol Sci* 277(1701):3735–3743.
33. Lewinsohn TM, Roslin T (2008) Four ways towards tropical herbivore megadiversity. *Ecol Lett* 11(4):398–416.
34. Connell JH (1971) On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. Dynamics of populations. *Proceedings of the Advanced Study Institute on Dynamics of Numbers in Populations*, eds den Boer PJ, Gradwell GR (Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands), pp 298–312.
35. Bagchi R, et al. (2014) Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* 506(7486):85–88.
36. Fox LR, Morrow PA (1981) Specialization: Species property or local phenomenon? *Science* 211(4485):887–893.
37. Bickford D, et al. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22(3):148–155.
38. Price PW (1980) *Evolutionary Biology of Parasites* (Princeton Univ Press, Princeton), pp xi–237.
39. Murcia C, et al. (2014) A critique of the 'novel ecosystem' concept. *Trends Ecol Evol* 29(10):548–553.
40. Tallamy DW (2004) Do alien plants reduce insect biomass? *Conserv Biol* 18(6): 1689–1692.
41. Bascompte J, Jordano P (2007) Plant-animal mutualistic networks: The architecture of biodiversity. *Annu Rev Ecol Evol Syst* 38(38):567–593.
42. Erdös P, Rényi A (1959) On random graphs. *Publ Math* 6:290–297.
43. Proulx SR, Promislow DE, Phillips PC (2005) Network thinking in ecology and evolution. *Trends Ecol Evol* 20(6):345–353.
44. Lewinsohn TM, Prado PI, Jordano P, Bascompte J, Olesen JM (2006) Structure in plant-animal interaction assemblages. *Oikos* 113(1):174–184.
45. Newman ME (2005) Power laws, Pareto distributions and Zipf's law. *Contemp Phys* 46(5):323–351.
46. Janz N, Nylin S (2008) The oscillation hypothesis of host-plant range and speciation. *Specialization, Speciation and Radiation: The Evolutionary Biology of Herbivorous Insects*, ed Tilmon KJ (Univ of California Press, Berkeley, CA), pp 203–215.
47. Courtney SP (1986) The ecology of pierid butterfies: Dynamics and interactions. *Adv Ecol Res* 15:51–131.
48. 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(9):2719–2724.
49. Arnold BC (2008) Pareto distribution. *Encyclopedia of Statistical Sciences*, 10.1002/0471667196.ess1910.pub3.
50. Aban IB, Meerschaert MM, Panorska AK (2006) Parameter estimation for the truncated Pareto distribution. *J Am Stat Assoc* 101(473):270–277.
51. Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. *SIAM Rev* 51(4):661–703.
52. D'Agostino RB (1986) *Goodness-of-Fit-Techniques* (CRC, Boca Raton, FL).
53. Zaninetti L, Ferraro M (2008) On the truncated Pareto distribution with applications. *Centr Eur J Phys* 6(1):1–6.
54. Deidda R (2010) A multiple threshold method for fitting the generalized Pareto distribution to rainfall time series. *Hydrol Earth Syst Sci* 14(12):2559–2575.
55. Schoenberg FP, Patel RD (2012) Comparison of Pareto and tapered Pareto distributions for environmental phenomena. *Eur Phys J Spec Top* 205(1):159–166.
56. Heywood VH (1993) *Flowering Plants of the World* (Mayflower Books, New York).

ECOLOGY

The global distribution of diet breadth in insect herbivores

Matthew L. Forister, Vojtech Novotny, Anna K. Panorska, Leontine Baje, Yves Basset, Philip T. Butterill, Lukas Cizek, Phyllis D. Coley, Francesca Dem, Ivone R. Diniz, Pavel Drozd, Mark Fox, Andrea Glassmire, Rebecca Hazen, Jan Hrcek, Joshua P. Jahner, Ondrej Kaman, Tomasz J. Kozubowski, Thomas A. Kursar, Owen T. Lewis, John Lill, Robert J. Marquis, Scott E. Miller, Helena C. Morais, Masashi Murakami, Herbert Nickel, Nick Pardikes, Robert E. Ricklefs, Michael S. Singer, Angela M. Smilanich, John O. Stireman, Santiago Villamarín-Cortez, Stepan Vodka, Martin Volf, David L. Wagner, Thomas Walla, George D. Weiblen, Lee A. Dyer

Supporting Information, Appendices *S1 – S6*

Appendix S1, Collection and curation of data

We used the most complete assemblage of site-specific plant-herbivore interaction data in existence, but a dataset of this size is characterized by inconsistencies and sampling artifacts that cannot be entirely mitigated. The data used in this study are comparable in that they are all based on collections and rearings of individual herbivores to document local plant-insect interactions. Some collections utilized identical methods across numerous sites, and these miscible subsets yield the same patterns that we report for the global data. For the full complement of sites, differences among datasets necessarily exist, and unique methods of collection and processing are described below for each dataset or collection of datasets (with additional details in Tables S1 and S2). In particular, datasets differ in the extent to which pre-analysis filters were applied. For example, at the first five sites described below (Lepidoptera from Arizona, Costa Rica, Ecuador, Great Basin and Louisiana), an insect has to have been found in association with a specific plant at least five times to be considered for analysis. For the majority of our other datasets (e.g. Lepidoptera from the Czech Republic and Papua New Guinea, and most of the guild-specific datasets), singletons were excluded (i.e. a plant-insect association that was observed only once is not analyzed). In a smaller number of additional cases, a quantitative filter was not applied (e.g. Japan, Belize), typically because the data were considered sufficiently reliable by the primary investigators, and some of these include already-published, stand-alone datasets (e.g., Canada).

Methods unique to each dataset, Lepidoptera

Arizona, Costa Rica, Ecuador, Great Basin, and Louisiana. Collecting at all five sites covered a broad range of latitude and longitude within the respective states and countries, including all of Arizona (USA; most collecting was at approximately 31°53'N, 109°12'W), all of Louisiana (USA; collecting was centered at 29°56'N, 90°7'W), all of Costa Rica (most collecting was at La Selva Biological Station, 10°26'N, 84°0'W), all of Ecuador (collecting centered at approximately 0°35'S, 77°53'W), and a large number of sites across the Sierra Nevada mountains and Great Basin desert (collecting centered at approximately 39°39'N, 119°46'W). Lepidopteran larvae at all sites were collected both opportunistically along trails, watercourses, and on undirected walks through the forests, and quantitatively using 10 m diameter plots as search areas. Plots were divided into four equal wedges, and one person spent 30 min looking for caterpillars on all the plants within each wedge. At all sites caterpillars were collected from all plant taxa and growth forms (herbs, vines, shrubs, trees) on which they were encountered.

All collected caterpillars were reared individually in clear plastic bags or glass jars in rearing facilities at ambient temperature and humidity. Fresh food in the form of new foliage from the same plant species from which the caterpillar was collected was placed in containers as needed. All pupae were checked daily to collect any adult Lepidoptera or parasitoids that emerged.

Voucher specimens of the focal plants and all first-time food plants were collected and pressed to insure accurate taxonomic identification and deposited at appropriate institutions. Initial identifications of insects were made by parataxonomists and then confirmed by taxonomic specialists. Voucher specimens of the insect species can be found at Tulane University, InBio (Costa Rica), the Museo National de Ciencias Naturales (Ecuador), and other collaborating institutions (see www.caterpillars.org for a list of participating institutions). Tachinid parasitoids reared from caterpillars at the Ecuador site were mounted, identified to genus (using Wood and Zumbado (1) and other resources) and sorted into morphospecies by JOS. When possible, specimens were identified to species with the aid of D.M. Wood and/or reference to specimens in the U.S. National Museum of Natural History and the Canadian National Collection of Insects (Ottawa, Canada). Some morphospecies were confirmed using mtDNA COI sequences. A pre-analysis filter of at least 5 observations (for any plant-insect interaction) was implemented for all of these datasets.

Table S1. Thirteen focal sites, with covariates used in analyses examining latitudinal gradients in specialization, including: latitude (decimal degrees), area and years sampled, the number of Lepidoptera species, and the number of rearing records studied; also the number of plant families and species studied at each site in the last column.

Site	Degrees lat.	Lep. species	Records	Area (ha)	Years	Plant fam. / sp.
Canada	47.25	671	131,431	80,000,000	20	27 / 257
Czech Republic	43.00	74	1,475	300	3	10 / 15
Japan	42.68	174	1,181	1000	2	18 / 51
Connecticut	41.50	116	744	1,437,100	26	45 / 175
Ohio	40.00	273	3,294	500,000	4	35 / 98
Great Basin	39.00	25	7,615	3,000,000	5	14 / 29
Arizona	32.21	186	14,229	3,000,000	18	47 / 176
Louisiana	31.00	328	14,505	500,000	12	67 / 196
Brazil	15.92	568	5,614	10,000	12	42 / 109
Costa Rica	10.42	1,321	35,103	2,400	22	81 / 454
Panama	9.17	401	4,536	1,600	10	50 / 210
Papua New Guinea	5.00	229	18,632	1,500	14	30 / 88
Ecuador	0.41	2,122	90,622	28,000,000	13	79 / 318

Brazil. We collected lepidopteran larvae in cerrado sensu stricto areas near Brasilia, DF, Brazil (15°45'S, 47°50'W) from 1991 to 2005. The region experiences a marked dry season from May through September (mean month precipitation of 24 mm). Caterpillars were collected opportunistically but on most of the plant species we collected once a week throughout a year. The plants examined included shrubs, trees and some herbs, varying in height from 0.50 to 2.50 m. All were censused during morning hours. The caterpillars were individually reared in plastic jars on leaves of the plant on which they were found, at ambient temperature and humidity. Leaves were replaced every two days, and any parasitoid emergence was recorded. Voucher specimens are housed at the Entomological Collection of Zoology Department, University of Brasilia, and were identified by Vitor Osmar Becker, Associate Researcher, Department of Zoology, University of Brasilia. A pre-analysis filter of 5 observations was implemented for the Brazil Lepidoptera data.

Canada. Collection records of individual caterpillars or groups of caterpillars that yielded adult Lepidoptera (moths) were entered from hand-written rearing records provided by the Canadian Forest Insect Survey (CFIS, now the Forest Insect Disease Survey), Sault Ste. Marie, Ontario, Canada. Larvae were collected from 1936-1955 by hand picking and beating foliage throughout the provinces of southern Canada (centered at approximately 53°N, 101°W). Collectors included forest biology rangers and associates of the CFIS (2). Collected larvae were reared at regional centers on foliage from the host plant on which they were collected. Identifications of larvae and reared adults were made by the Insect Systematics Unit of the Survey. A combination of published data from these surveys (3-6) and the original hand-written rearing records were used to construct the caterpillar-host plant database. Plant and insect nomenclature were updated as needed to reflect the current state of taxonomic understanding.

Connecticut. Caterpillars (Lepidoptera) were collected both opportunistically (1988-2013) across many communities throughout the state of Connecticut and in adjoining areas of Massachusetts and Rhode Island (centered at approximately 41°33'N, 72°39'W), as well as quantitatively using haphazardly chosen branches from focal tree taxa at three sites in Middlesex County, Connecticut (2004-2012). For quantitative sampling, collection effort on each of eight tree taxa (*Quercus rubra* group, *Quercus alba*, *Fagus grandifolia*, *Carya* spp. (Eucarya group), *Prunus serotina*, *Acer rubrum*, *Betula lenta*, and *Hamamelis virginiana*) was equalized per collection day. Saplings and low branches of larger trees were sampled (ground level-3 m in height). Much of the opportunistic and all of the quantitative collecting occurred during the peak caterpillar season in Connecticut (May- July), although limited collecting efforts extended throughout the growing season (May-October).

Caterpillars were reared individually in vials, plastic cups, bags, or boxes in the laboratory under ambient conditions. Caterpillars were fed foliage from the plant taxon upon which they were collected. Reared adults were spread to facilitate taxonomic identification. Identifications were based on reared adult specimens, or, for relatively distinct and well-described species, on larval features using Wagner (7) and Wagner et al. (8-10), or both. Voucher specimens are housed at the University of Connecticut (Storrs) and Wesleyan University.

To correct for possible bias due to the combination of opportunistic and restricted quantitative sampling approaches, we limited the host records used here to a randomly chosen set of 116 species of Zygaenoidea, Papilionoidea, and macrolepidoterans meeting certain criteria.

116 The random set of species was drawn from a Connecticut state checklist of Lepidoptera (2410
117 species) compiled by DLW. We used a random number generator to choose candidate species
118 from the numbered checklist. Criteria for inclusion were: 1) externally feeding larvae (including
119 leaf-tiers), 2) host range including at least some woody plants, and 3) the availability of and our
120 confidence in rearing records of larvae on natural, reliably identified host plants. If a randomly
121 chosen species met these criteria, it was included. If not, it was discarded and a new randomly
122 chosen species was considered. This iterative process continued until the data set included 116
123 species, which we estimated (based on previous experience with similar data) was a sufficiently
124 large sample size for analysis.

125

126 **Czech Republic.** We collected all externally feeding, leaf-tying and rolling caterpillars
127 (Lepidoptera) from accessible foliage of 15 focal, locally-common woody plant species (14
128 species listed in Novotny et al. (*11*) and *Acer pseudoplatanus*) in the Poodri Protected Area
129 (18°03-13'E, 49°42-48'N, 200 m asl., Czech Republic). The study area of 300 ha included three
130 fragments of the primary floodplain forest along a 20 km long section of the Odra River. The
131 forest vegetation was dominated by *Quercus*, *Ulmus*, *Tilia*, *Prunus* and *Fraxinus*. The study
132 plant species represented 85 ± 2.4% of the total forest basal area according to 62 plots of 25 x 25
133 m each, where all plants ≥ 5cm in diameter at breast height (DBH) were recorded. The annual
134 average temperature was 7-8.5°C, the annual average rainfall 600-800 mm. Insect sampling
135 effort amounted to 150 m^2 of foliage inspected per tree species. Each caterpillar was provided
136 with fresh leaves of the plant species from which it was collected and only those that fed were
137 retained in the analyses. Larvae were identified to morphospecies and/or reared to adults. All
138 insects assigned to morphospecies were later verified and identified by taxonomic specialists.
139 Vouchers are deposited at the University of Ostrava, Ostrava. The pre-analysis filter of excluding
140 singleton observations was implemented for the Czech Republic Lepidoptera.

141

142 **Ohio**. We sampled caterpillars (Lepidoptera) from temperate deciduous forest fragments in
143 Southwestern Ohio (ca. 39ºN, 84ºW) from 2006-2009. All woody plant feeding caterpillars, as
144 well as some herbaceous feeding species, were collected along 100m transects from 19 forest
145 fragments ranging in size from 6 to 800 ha. Each fragment was sampled with between 2 and 20
146 transects. All caterpillars within one meter on each side of the transect line, from ground level to
147 a height of ca. 2.5 m, were recorded and collected. Caterpillars were placed in plastic bags with
148 foliage from their host plants and later transferred to plastic tubs placed in an environmental
149 chamber with temperatures and light regimes mimicking the seasonal temperatures and light
150 regimes of the region. Every other day they were fed new leaves of plant species on which they
151 were found until they died, pupated, or a parasitoid emerged. When they were near pupation,
152 plant material was replaced with moistened peat moss in which they could pupate.
153   All caterpillars collected were identified to the lowest taxonomic level possible based on
154 morphological appearance, distribution, host plant use, and seasonality using Wagner (7). Once
155 adults emerged, specimens were mounted and identified with the use of Covell (12), Wagner (7),
156 Microleps.org, The North American Moth Photographer's Group
157 (http://mothphotographersgroup.msstate.edu), and other traditional and digital resources. Some
158 specimens were taken to the Ohio State University insect collection for comparison and
159 identification. Specimens that could not be identified retained morphospecies designations.
160 Vouchers of all taxa were deposited in the Wright State University insect collection. Insects that
161 were only observed once from a single plant were excluded from analyses.

162

Panama. We collected caterpillars (Lepidoptera) on Barro Colorado Island (BCI), Panama (9ºN, 80ºW) from 1996-2005 with some additional collections in 2013. The island is maintained and protected by the Smithsonian Tropical Research Institute and is part of a larger forested corridor that extends from the Atlantic to Pacific coasts. BCI experiences a marked dry season that is usually four months long, and the vegetation is classified as tropical, moist forest (13-15).

Caterpillars were collected opportunistically from the island's diverse understory of shade-tolerant plants representing a variety of life histories. The plant species included shrubs, juvenile lianas and immature trees with growth strategies that differed widely, even within plant genera. We reared all caterpillars individually in closed plastic containers or plastic bags at ambient temperature in a screened and shaded porch. We fed them leaves of the same species and age as those on which they were initially found. Leaves were replaced with fresh ones at least every other day. We photographed and reared specimens and entered them into our database (http://www.biology.utah.edu/~coley/database.htm). Using larvae and adult specimens, we identified the lepidopterans to the lowest taxon possible. Voucher specimens are stored on BCI, and some duplicate specimens are with experts for identification. Host plants were identified to species using Croat (14) and by comparison with herbarium specimens.

Papua New Guinea. All externally feeding caterpillars (Lepidoptera), including leaf-tiers and rollers, were sampled from 88 woody species of plants, representing all major lineages of flowering plants, at three study sites (Baitabag, Ohu, and Mis Villages; 145°41-8' E, 5°08-14' S, 0-200 m asl., Papua New Guinea) within a 10 x 20 km area, encompassing a mosaic of secondary and primary lowland hill forest. The annual average temperature was 26.5°C, and the annual average rainfall 3,600 mm. The sampling took place within approximately 1,500 ha of primary and secondary forests. Each tree species was sampled for at least one year between 1994-2008. The sampling effort amounted to 1,500 m^2 of foliage per tree species, obtained from multiple conspecific individual trees. All caterpillars were provided with fresh leaves of the plant species from which it was collected and only those that fed were retained in the analyses. Larvae were identified to morphospecies; adults were identified by genitalia, DNA barcoding and consultation with taxonomic specialists. Vouchers are deposited at the National Agricultural Research Institute of Papua New Guinea and Smithsonian Institution, Washington. See Novotny et al. (16, 17, 11) for more information. Any singleton observations were excluded from the Papua New Guinea Lepidoptera data.

Parasitoids were reared from the above caterpillar sampling encompassing 38 tree species. Reared parasitoids were linked with their host through caterpillar morphospecies. Parasitoids were mounted, morphotyped by JH and identified by taxonomists listed in Hrcek et al. (18). A selection of the parasitoid specimens was DNA barcoded and any identifications in conflict with DNA barcodes were re-examined. Parasitoids belonged to Hymenoptera: Braconidae, Ichneumonidae, Chalcidoidea and Bethylidae, and Diptera: Tachinidae; see Hrcek et al. (18) for more information. Singleton observations were excluded prior to analysis.

Japan. We collected caterpillar (Lepidoptera) at the Tomakomai Research Station in Hokkaido, Japan (42° 42'N, 141° 36'E) from 2008 to 2009. This cool-temperate mixed forest receives 1,161 mm of annual precipitation, and the average annual temperature is 5.6°C. Maple (*Acer mono*), linden (*Tilia japonica*), and oak (*Quercus crispula*) dominate the forest. The canopy ranges from 15 m to 25 m in height. Deciduous trees break bud in early to mid-May and shed their leaves in

late October. Caterpillars were collected using truck-mounted elevated work platforms (cherry-pickers) on 51 plant species representing 18 families including four conifer species. Samples were taken twice (spring and summer) for each year, in total four times during the survey. For each sampling occasion, three branches from three tree individuals, i.e. 9 branches in total, were sampled (5.1 ± 3.1, mean $\pm$ sd cm in diameter) and all the caterpillars were picked up by hand. The caterpillars were individually reared in plastic cups on leaves of the plant on which they were found at ambient temperature and humidity. Voucher specimens are housed at the Chiba University.

Methods unique to each dataset, all herbivores

Guilds studied in Papua New Guinea. Seven guilds (listed below, classified as in Novotny et al. (19), with some modifications) were studied near the villages of Baitabag, Mis and Ohu near the town of Madang (Papua New Guinea), within a 20×10 km area comprising a successional mosaic of disturbed and mature lowland rainforest (5°08'-14'S, 145°7'-41'E, 50–200 m above sea level, Madang Province). The vegetation has been classified as mixed evergreen rain forest on Latosol (16, 17, 19) with a humid climate (mean annual rainfall 3600 mm), a mild dry season from July to September, and mean annual temperature of 26°C. All trophic interactions were confirmed by feeding experiments for adults or rearing for larvae. Plant-herbivore trophic interactions supported by singletons were excluded from the analysis. Plant vouchers are deposited at the PNG National Herbarium (Forestry Research Institute, Lae), insect vouchers at the Smithsonian Institution (USA) and the Institute of Entomology of the Academy of Sciences (Ceske Budejovice, Czech Republic).

Adult leaf chewers, Papua New Guinea. All externally feeding adults (Orthoptera, Phasmatodea and Coleoptera) were sampled in Madang from 59 native rainforest woody species representing 19 families. Insects were hand-collected from 1500 m^2 of foliage per plant species over the period of approximately 12 months, sampling young and mature foliage from multiple individual trees. The sampling took place from 1995 to 2002 (16, 23, 24). All individuals were tested in a no-choice feeding experiment on the leaves of the plant species they were collected from; only feeding individuals were included in the analysis. As described above, singleton observations were excluded.

Larval leaf chewers, Papua New Guinea. All externally feeding, leaf rolling and leaf tying holometabolous larvae (Lepidoptera and Coleoptera) were sampled in Madang from 88 woody species representing 31 plant families. Insects were hand-collected from 1500 m^2 of foliage per plant species over the period of approximately 12 months, and reared to adults as far as possible. The sampling continued from 1995 to 2008 (15, 17, 23, 24). As described above, singleton observations were excluded.

Leaf miners, Papua New Guinea. All leaf-mining larvae (Lepidoptera, Coleoptera, Diptera) were sampled in Madang from 76 woody species representing 31 plant families. Insects were hand-collected from 1500 m^2 of foliage per plant species over the period of approximately 12 months, and reared to adults as far as possible. The sampling continued from 2006 to 2008 (19, 23). Only reared adults were analyzed. As described above, singleton observations were excluded.

254

255

Table S2. Datasets of herbivore feeding guilds, with details relevant to pairwise comparisons between tropical and temperate communities, including the numbers of plant families and species associated with insects used in analyses, as well as the numbers of herbivore species. For other details, see supplementary text, and note that plant species data were not available for two datasets (leaf and phloem suckers from Germany). The sites listed below are only partially overlapping with the sites used for analyses involving only Lepidoptera; see Table S1 and Fig. 1*A*.

Site	Feeding guild	Major herbivore taxa	Herbivore species	Plant fam. / sp.
Tropical datasets				
Papua New Guinea	Adult leaf chewers	Orthoptera, Phasmatodea, Coleoptera	245	19 / 59
Papua New Guinea	Larval leaf chewers	Lepidoptera, Coleoptera	326	30 / 86
Papua New Guinea	Leaf miners	Lepidoptera, Coleoptera, Diptera	124	28 / 63
Belize	Leaf miners	Lepidoptera, Coleoptera, Diptera	81	20 / 60
Papua New Guinea	Leaf suckers	Auchenorrhyncha: Typhlocybinae	36	13 / 33
Papua New Guinea	Phloem suckers	Auchenorrhyncha	29	8 / 14
Papua New Guinea	Gallers	Diptera, Coleoptera, Lepidoptera, Hymenoptera	76	13 / 27
Papua New Guinea	Xylem chewers	Cerambycidae	40	14 / 23
Panama	Xylem chewers	Cerambycidae	22	6 / 9
Temperate datasets				
Czech Republic	Adult leaf chewers	Coleoptera	20	9 / 14
Czech Republic	Larval leaf chewers	Lepidoptera, Hymenoptera, Coleoptera	94	10 / 15
Great Britain	Leaf miners	Lepidoptera, Coleoptera, Diptera	31	11 / 20
Poland	Leaf miners	Lepidoptera, Coleoptera, Diptera	34	10 / 17
Great Britain	Leaf suckers	Auchenorrhyncha: Typhlocybinae	55	10 / 32
Germany	Leaf suckers	Auchenorrhyncha	81	12 / -
Germany	Phloem suckers	Auchenorrhyncha	84	17 / -
Czech Republic	Gallers	Diptera, Hymenoptera, Sternorrhyncha	43	7 / 9
Czech Republic	Xylem chewers	Cerambycidae	46	11 / 22

256

257 **Leaf suckers, Papua New Guinea.** Mesophyll-sucking larvae (Auchenorrhyncha, Cicadellidae:
258 Typhlocybinae) were sampled in Madang from 56 woody species representing 22 plant families.
259 Insects were hand-collected from 1500 m^2 of foliage per plant species over the period of six
260 months and reared to adults (26). Only reared adult males were used in the analysis; in
261 Typhlocybinae, the taxonomy is based mostly on the morphology of male genitalia whereas

females are often impossible to identify morphologically. As described above, singleton observations were excluded.

Phloem suckers, Papua New Guinea. Phloem-sucking larvae of leafhoppers and planthoppers (Auchenorrhyncha) were sampled in Madang from 14 woody species representing eight plant families. Insects were hand-collected from 1500 m^2 of foliage per plant species from 2007 to 2009 and reared to adults on live saplings caged in forest understorey (27). Only reared adults were used in the analysis. As described above, singleton observations were excluded.

Gallers, Papua New Guinea. Insect-induced galls were collected from 32 species of woody plants, representing all major lineages of flowering plants in Madang, from August 2010 to March 2011. Galls were hand-collected from the accessible foliage of multiple conspecific individuals of each species, representing a total sampling effort of approximately 69 hours of searching for galls on each tree species. Harvested galls were initially morphotyped according to gall morphology within host plant and reared in clear plastic bags for a period of one month in order to obtain adults. In addition, immature insect stages were obtained from gall dissections. Reared insects were identified to morphospecies, and to species where possible by taxonomic specialists, and both adults and/or immature insects of as many morphospecies as possible were DNA barcoded. Galling insects were used in analyses of species-level diet breadth, but not family-level diet breadth because all insect species were specialists with a diet breadth of one host plant family, from which a distribution of diet breadth could not be estimated. As described above, singleton observations were excluded.

Xylem chewers, Papua New Guinea. Wood-boring beetles (Cerambycidae) were reared from standardized timber baits of freshly cut wood from 26 tree species representing 15 families in Madang from 2000 to 2003. Eight trunk sections, each weighing 20 kg, per tree species were sawn from freshly felled trees and exposed to ovipositing females in the forest canopy and understorey. After three weeks each section was individually enclosed in a rearing cage and emerging insects collected for six months (19). As described above, singleton observations were excluded.

Leaf miners, Belize. All leaf-miners (Coleoptera, Diptera, Lepidoptera) were sampled from low vegetation along trails in deciduous seasonal and deciduous semi-evergreen seasonal forest near the Las Cuevas Research Station (Chiquibul Forest Reserve) in south-west Belize. All vegetation up to a height of 2 m was sampled within 34 plots 100 x 2.5 m and miners reared to adults as far as possible. Mining species were morphotyped using the morphology of mines and adults. The sampling continued for one year, from 1997 to 1998. Overall, 70 plant species from 25 families were sampled in proportion to their abundance in the study plots (25).

Xylem chewers, Panama. Cerambycid beetles were reared from standardized timber baits exposed in canopy and understorey of lowland wet forests in San Lorenzo Protected Area (9°16'56"N, 79°58'26"W; 150-180 m asl, Colón Province; mean annual temperature 26.0ºC, mean annual rainfall 3,139 mm), then caged for rearing. 12 tree species from nine families were sampled using 8 baits with combined weight ca 100 kg of wood per tree species. Singleton observations of plant-insect interactions were removed.

Adult leaf chewers, Czech Republic. All adult herbivores (Coleoptera) were sampled from accessible foliage of 15 focal, locally common woody plant species in the Poodri Protected Area (49°42-48'N, 18°03-13'E; 200 m above seas level, mean annual temperature 7-8.5°C, mean annual rainfall 600-800 mm). The study area of 300 ha included three fragments of the primary floodplain forest dominated by *Quercus*, *Ulmus*, *Tilia*, *Prunus* and *Fraxinus*. The study plant species represented ~85% of the total forest basal area of trees in the forest. Insect sampling continued throughout the growing seasons during 1999-2001. Sampling effort amounted to 150 m^2 of foliage inspected per tree species. Each herbivore was tested in a no-choice experiment on the leaves of the plant it was collected from. Only feeding individuals were retained in the analysis. Plant-herbivore trophic interactions supported by singletons were excluded from the analysis (11).

Larval leaf chewers, Czech Republic. All externally feeding, leaf-tying and rolling larvae (Lepidoptera, Hymenoptera, Coleoptera) were sampled from the same trees as mobile chewers (see above). Larvae were identified to morphospecies and/or reared to adults. All insects assigned to morphospecies were later verified and identified by taxonomic specialists. Plant-herbivore trophic interactions supported by singletons were excluded from the analysis (11).

Leaf miners, Great Britain. Leaf-miners were sampled from 20 herbaceous species and 11 families in a damp semi-natural grassland 1.8 ha in size (Rush Meadow) at Silwood Park, Berkshire, UK (51°24'48"N, 0°39'02"W) from May to October 1999. Leaves containing living miners were placed in sealed plastic tubs to rear hosts. Diapausing hosts were over-wintered at ambient temperatures in a well-ventilated outdoor shed. The trophic interactions recorded are based on unstandardized collections made in a stratified manner throughout the site, and reflect only sampled hosts that were reared to generate identifiable adult leaf-miners.

Leaf miners, Poland. The data on leaf-mining insects of the Białowieża National Park, Poland (52°45'N, 23°53'E; 145-175 m above sea level, mean annual temperature 6.8°C, mean annual rainfall 641 mm) were published by Michalska (28). They include all leaf-miners and their host plants recorded in the national park, i.e. 152 km^2 of predominantly old growth-forest (forest associations *Circaeo-Alnetum, Carici elongatae-Alnetum, Peucedano-Pinetum, Querceto Carpinetum stachyetosum* and *Pineto-Quercetum serratuletosum*) during the years 1967-1972. The present analysis includes only leaf miners on woody plant species.

Leaf suckers, Great Britain. Feeding records for mesophyll-sucking leafhoppers (Auchenorrhyncha, Typhlocybinae) were collected from woody host species at 22 study sites in Great Britain and confirmed by rearing of larvae (29). Trophic interactions supported by at least 5 rearing records were included in the analysis.

Leaf and phloem suckers, Germany. Mesophyll cell and phloem sucking Auchenorrhyncha were sampled within the 1,117 km^2 area of the administrative district of Göttingen, Germany (51°17-38'N, 9°32'-10°22'E, 120–530 m above sea level, mean annual temperature 8.7°C, mean annual rainfall 645 mm) for 25 years (1990 – 2014). Insects were collected from woody plant species mainly between May and October through sweep-netting and direct search on the host, without any standardized design but aiming at extensively covering all potential host species. The study area is a semi-open landscape with fields and managed forests, as well as numerous

354 hedgerows, alleys and single trees. Singletons were removed prior to analyses; data are detailed
355 in Nickel (30, 31).
356
357 **Gallers, Czech Republic.** All gallers were as far as possible completely sampled from plants
358 with DBH>5cm at two 0.1 ha plots in broadleaf floodplain forest in Southern Moravia, Czech
359 Republic (48°48'N, 17°5'E, 152 asl, and N48°41′, E16°56′, 164 m above sea level, mean annual
360 temperature 9°C, mean annual rainfall 525 mm). The terrain was flat and the dominant trees
361 included species of *Quercus, Fraxinus, Carpinus, Acer* and *Populus*. Galls were collected from
362 83 plant individuals representing nine tree species. The total sampling effort amounted to 12,500
363 m^2 of foliage. The sampling was carried out during the 2013 vegetation season. Canopies were
364 accessed by tree-felling at one of the 0.1 ha plots and from an elevated work platform (cherry-
365 picker) at the second one. In the felled plot, sampling was carried out during the peak of miner
366 abundance (mid May – mid June), and in the cherry-picker plot samples were taken throughout
367 the whole season from May to August. Galls were morphotyped and reared to adults for further
368 identification by specialists. Note that galling insects were used in analyses involving species-
369 level diet breadth (e.g. Fig 3), but were not used in analyses involving family-level diet breadth
370 because the temperate gallers (as well as the tropical gallers, from Papua New Guinea, see
371 above) all had a family-level diet breadth of 1 (host family), from which it was not meaningful to
372 estimate a diet breadth distribution. Singleton observations of interactions were excluded.
373
374 **Xylem chewers, Czech Republic.** Cerambycid beetles were studied in the lowland and foothill
375 forests along the Dyje river in the Podyji National Park and the Lower Moravia UNESCO
376 Biosphere Reserve (N48°46-51', E15°50'-16°50'; 160-350 m above sea level; mean annual
377 temperature 9 °C, mean annual rainfall 524 mm). The dominant trees include *Quercus* spp.,
378 *Fraxinus* spp., *Carpinus betulus*, and *Acer* spp. Historically, the forests had been managed as
379 coppice, coppice with standards or pasture woodland. The cerambycids were reared from
380 standardized timber baits of 22 woody plant species representing 11 families, including two
381 conifers. Each plant species was sampled using 12 baits with average weight of 15 kg each,
382 exposed in canopy and understorey and subsequently caged for insect rearing. Singleton
383 observations of interactions were excluded.
384
385

385 *Appendix S2, Taxonomic and phylogenetic diet breadth*

386

387 Most of the analyses that we report utilized both family and species-level diet breadth, and
388 results obtained with the two indices were always similar, which is perhaps not surprising as
389 family and species-level diet breadth are highly correlated, as can be seen in Fig. S1. Examples
390 of the distribution of both species and family-level diet breadth are shown in Fig. S2 for
391 Lepidoptera; and an example of family-level diet breadth is shown in Fig. S3 for parasitoids.

392 The shape parameter (α) from the truncated, discrete Pareto distribution fit to family-level
393 diet breadth changes with latitude, as can be seen in Fig. 2*A* (main text). The relationship
394 between α and latitude is significant when modeled with the number of rearing records as a
395 covariate proxy for sampling effort: $F_{2,10} = 8.87$, $R^2 = 0.64$, $P = 0.0061$; partial regression
396 coefficient for latitude = -0.023, $P =$
397 0.0046; coefficient for rearing records
398 = 0.062, $P = 0.69$. Results are
399 essentially identical with different
400 covariates: the area of sites, the
401 number of years that records have been
402 taken at each site, the number of
403 records, or the number of herbivore
404 species (years, records, area and
405 number of species are not individually
406 significant, but the overall model and
407 the effect of latitude are always
408 significant; in all cases covariates
409 conform to assumptions of standard
410 linear models). Although α changes
411 with latitude, the goodness of fit of the
412 Pareto distribution does not change
413 with latitude, as can be seen for
414 family-level diet breadth in Fig. S4*A*.

415 The shape of the discrete,
416 truncated Pareto for species-level diet
417 breadth also changes with latitude
418 (Fig. S4*B*): the associated partial
419 regression coefficient for latitude is -
420 0.015 ($P = 0.035$); results for the
421 whole model as follows: $F_{2,10} = 3.65$,
422 $R^2 = 0.42$, $P = 0.065$.

423 Phylogenetic diet breadth (PD)
424 similarly changes with latitude (Fig.
425 2*C*, main text). Using the number of
426 rearing records from each site as a
427 covariate for sampling effort, results
428 from the full model relating latitude to
429 PD are as follows: $F_{2,10} = 11.58$, $R^2 =$
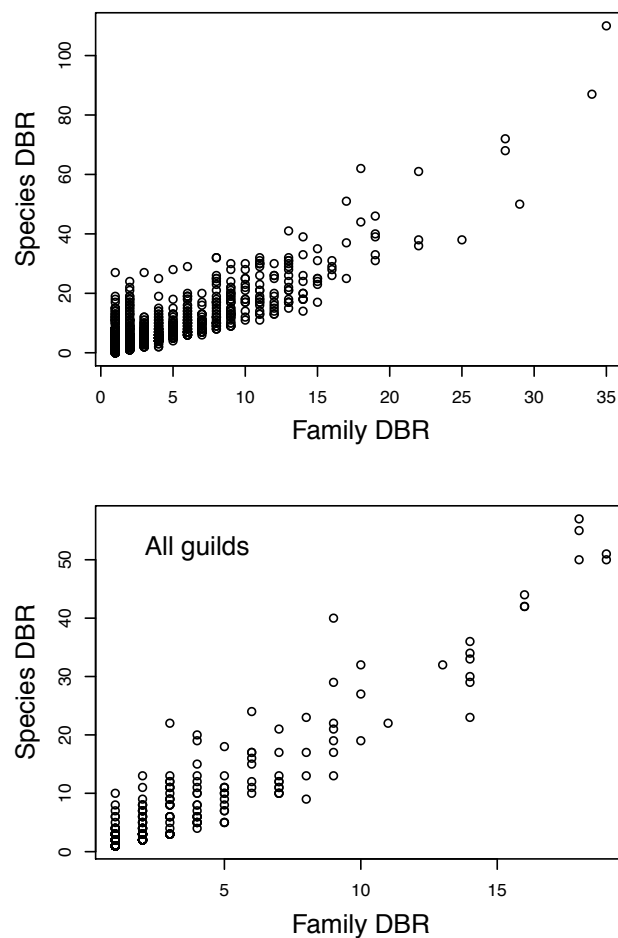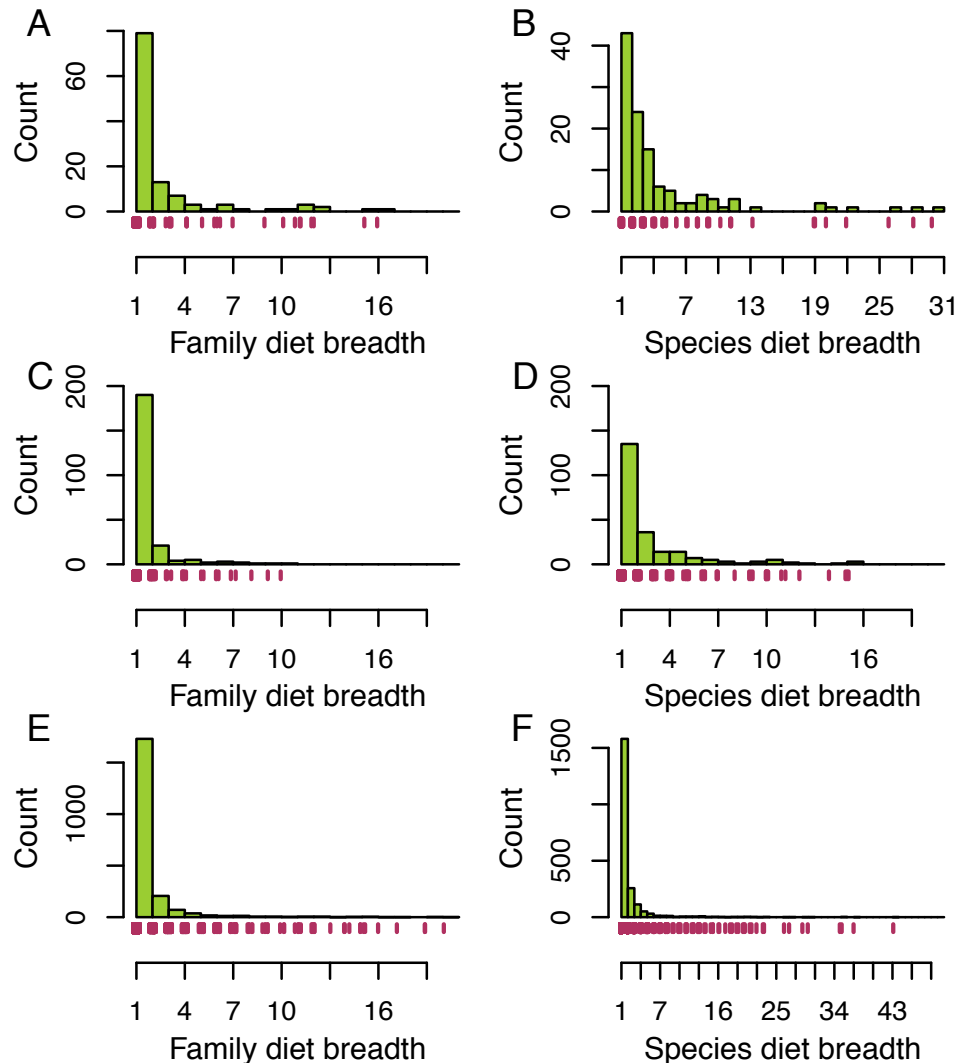430 0.70, $P = 0.0025$; for the latitude term:

Fig. S1. Species–level diet breadth against family–level diet breadth for the two major subset of the data: all Lepidoptera, and all guilds (excluding Lepidoptera). Each point illustrates species and family–level diet breadth for a single herbivore species.

2.35 partial regression coefficient, P < 0.001; for the sampling records term: 22.40 coefficient, *P* = 0.088.

Finally, both species-level and family-level diet breadth were investigated for taxonomic and geographic subsets of the data. Results for analyses of the Pareto distribution for subsets of the data are shown in Table S3, including focal sites for Lepidoptera, and for species within the top ten most frequently occurring families of Lepidoptera in our data. Also shown in Table S3 are the Pareto statistics for different herbivore guilds and for the diet breadth of Lepidoptera associated with the top ten most frequently-studied plant families.

Finally, fit is shown for parasitoids from one New and one Old World site.

For Pareto fit associated with the top ten most frequently studied plant families (Table S3), the species count is the number of associated herbivores not the number of plant species. With respect to the Pareto fit to parasitoid diet breadth (the last two rows), only family-level records were available. Included in Table S3 is a test statistic from a χ^2 test for goodness of fit and associated *P* values.

Fig. S2. Examples of the distribution of family and species–level diet breadth from three sites, from high to low latitude: Connecticut (A and B), Papua New Guinea (C and D), and Ecuador (E and F). Tick marks under plots mark individual observations for ease of visualization in the thin tail of the distribution. Note that the y axes differ in scale among plots as the number of Lepidoptera species differs among sites.

477
478
479
480
481
482
483
484
485
486
487
488
489
491
493
495
497
499
501
503
505
507
509
511
513
515
517
519
521
523
525
526
527
528
529
530
531
532
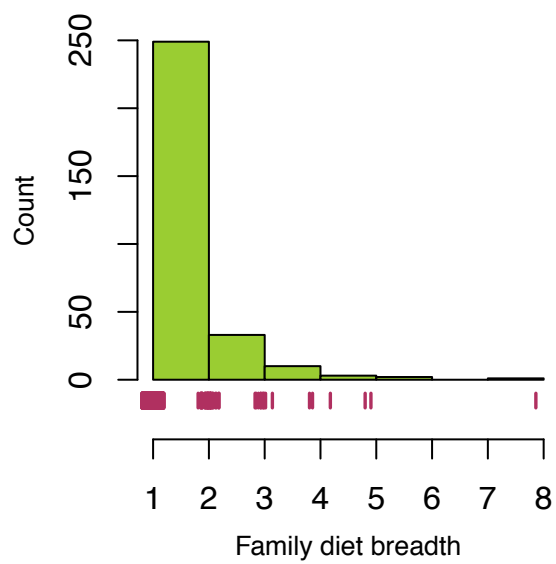
Fig. S3. Distribution of diet breadth for tachinid
flies, Ecuador. These are parasitoids that attack caterpillars.

Table S3. Pareto statistics for species-level and family-level diet breadth (DBR): alpha (the shape parameter) and the upper truncation parameter (the maximum observation) are shown, as well as the goodness of fit test statistic and associated *P* value (values less than 0.05 reject the fit of the Pareto distribution). Also shown is richness ("Sp.") for each site or taxon. For focal sites, percentages after site names are the percentages of species found in association with a single host species and a single host family as calculated from raw data (first and second values shown, respectively). The same values are reported for feeding guilds.

Site or taxon	Sp.	species-level DBR				family-level DBR			
		α	max.	test stat.	*P*	α	max.	test stat.	*P*
All Lepidoptera									
World	6,388	1.53	110	57.60	0.06	1.85	35	6.37	1.00
New world	5,565	1.50	110	47.78	0.19	1.82	35	6.90	1.00
Old world	679	1.11	31	3.70	1.00	1.45	16	13.29	0.58
Focal sites									
Canada (22%, 43%)	671	0.90	46	93.52	<0.01	1.32	15	18.42	0.19
Japan (63%, 67%)	174	1.10	14	4.50	0.98	0.98	7	2.13	0.91
Czech Republic (51%, 58%)	74	1.03	13	9.72	0.96	1.18	9	5.11	0.75
Connecticut (37%, 68%)	116	0.58	30	3.06	1.00	0.86	16	9.26	0.86
Ohio (45%, 58%)	173	1.02	26	5.76	0.95	0.99	15	3.42	1.00
Great Basin (56%, 60%)	25	0.29	9	6.19	0.62	0.49	9	2.29	0.97
Arizona (53%, 64%)	186	1.11	53	4.01	0.98	1.31	24	2.25	0.97
Louisiana (53%, 60%)	328	1.07	67	1.41	1.00	1.24	36	3.39	1.00
Brazil (67%, 77%)	568	1.33	33	3.23	1.00	1.46	19	16.44	0.56
La Selva (71%, 81%)	1321	1.62	61	20.47	0.77	1.95	28	6.36	0.96
Panama (78%, 92%)	400	1.72	18	0.78	1.00	2.39	10	9.35	0.41
Papua New Guinea (59%, 83%)	229	0.88	15	4.92	0.99	1.69	10	1.19	1.00
Ecuador (74%, 82%)	2122	1.52	43	8.05	1.00	1.79	20	5.99	1.00
Ten most frequently-studied Lepidoptera families									
Geometridae	1293	1.25	41	14.78	1.00	1.67	17	3.07	1.00
Noctuidae	826	1.23	45	23.09	0.88	1.67	19	6.22	0.99
Erebidae	419	1.03	109	18.90	0.59	1.28	35	5.63	1.00
Pyralidae	608	1.69	25	2.27	1.00	2.00	17	3.13	0.99
Tortricidae	256	0.70	33	4.26	1.00	1.20	19	1.78	1.00
Saturniidae	189	0.71	40	10.70	0.91	0.86	22	9.65	0.79
Nymphalidae	363	1.44	20	1.87	1.00	2.60	7	4.61	0.60
Notodontidae	328	1.16	33	3.13	1.00	1.56	11	1.08	0.98
Hesperiidae	317	1.66	27	2.82	0.99	2.31	5	0.56	0.97
Lymantriidae	47	0.48	60	12.61	0.32	0.66	22	11.49	0.18
Ten most frequently-studied plant families (as Lepidoptera hosts)									
Pinaceae	291	0.52	23	8.96	0.99	-	-	-	-
Betulaceae	484	1.52	11	5.27	0.88	-	-	-	-
Fabaceae	672	2.06	18	2.86	1.00	-	-	-	-
Salicaceae	455	1.77	9	5.74	0.68	-	-	-	-
Piperaceae	513	1.79	27	2.55	1.00	-	-	-	-
Asteraceae	448	1.66	16	5.83	1.00	-	-	-	-
Rosaceae	403	1.52	9	0.98	1.00	-	-	-	-
Fagaceae	320	1.56	13	4.01	0.95	-	-	-	-
Sapindaceae	356	1.90	7	2.69	0.85	-	-	-	-
Rubiaceae	384	1.87	8	0.71	1.00	-	-	-	-

Guilds

Leaf miners, Belize (89%, 99%)	81	2.08	8	1.20	0.88	6.22	2	7.88×10^{-13}	1.00
Leaf miners, Papua New Guinea (89% 99%)	124	2.03	9	1.85	0.93	4.96	3	3.13	0.21
Leaf miners, UK (77%, 97%)	31	1.96	5	0.52	0.77	4.73	2	4.82×10^{-13}	1.00
Leaf miners, Poland (82%, 97%)	34	2.53	3	0.094	0.95	3.44	3	2.08	0.35
Leaf suckers, Papua New Guinea (83%, 94%)	36	2.60	3	0.079	0.96	3.17	3	0.68	0.71
Leaf suckers, UK (59%, 76%)	55	0.58	17	2.03	0.96	1.32	7	2.82	0.83
Leaf suckers, Germany (NA, 70%)	81	-	-	-	-	1.31	7	2.20	0.90
Larval leaf chewers, Papua New Guinea (61%, 83%)	326	1.07	25	8.77	0.98	1.57	15	12.68	0.55
Larval leaf chewers, Czech Republic (56%, 66%)	94	1.19	13	1.60	1.00	1.39	9	2.48	0.96
Xylem chewers, Papua New Guinea (33%, 35%)	40	0.45	16	5.69	0.93	0.60	10	4.21	0.90
Xylem chewers, Panama (64%, 77%)	22	1.08	6	1.68	0.89	1.43	3	0.096	0.95
Xylem chewers, Czech Republic (35%, 48%)	46	0.69	19	20.89	0.035	0.91	9	4.69	0.79
Phloem suckers, Papua New Guinea (45%, 45%)	29	-	-	-	-	0.20	8	1.24	0.99
Phloem suckers, Germany (NA, 70%)	84	-	-	-	-	1.33	12	3.42	0.91
Adult leaf chewers, Papua New Guinea (43%, 55%)	245	0.46	57	5.02	1.00	0.68	19	8.63	0.97
Adult leaf chewers, Czech Republic (55%, 55%)	20	0.14	13	9.42	0.67	0.31	8	0.78	0.94
Gallers, Papua New Guinea (99%, 100%)	76	6.12	2	3.91×10^{-14}	1.00	-	-	-	-
Gallers, Czech Republic (86%, 100%)	43	2.22	2	1.74×10^{-15}	1.00	-	-	-	-

Parasitoids

Ecuador, Tachinidae	309	-	-	-	-	2.56	8	4.44	0.73
Papua New Guinea, wasps and flies	58	-	-	-	-	3.49	2	9.98×10^{-14}	1.00

534
535
536
537

Appendix S3, Change among sites in the distribution of diet breadth

The shape parameter (α) of the discrete, truncated Pareto distribution is a useful summary statistic for investigating change in the distribution of diet breadth (e.g. Fig. 2*A*, main text), but change in diet breadth can also be visualized with the maximum observation (the upper truncation parameter from the Pareto distribution) and quantiles (as in Fig. 2*B*). In Fig. S5, we show examples of the distribution of family-level diet breadth at a subset of sites in order to visualize the behavior of summary statistics including quantiles.

      Specifically, five parameters are shown in Fig. S5: α (the shape parameter from the discrete, truncated Pareto), β (the upper truncation parameter), and a selection of quantiles (the 99th, the 95th, and the 90th). Quantiles are a useful way to measure change in density throughout the tail as the frequency of relatively-specialized herbivores increases at lower latitudes. In the main text and Fig. 2*B*, change in the 90th quantile is

Fig. S4. (A) Goodness of fit of the Pareto distribution to family–level diet breadth (DBR) against latitude, with standard errors from 1000 bootstrap resamples. (B) Latitudinal trend in the shape parameter for species–level diet breadth (DBR). Greater values of the shape parameter indicate distributions of diet breadth at individual sites that include a greater portion of relatively specialized herbivores. Standard errors for individual sites (points) are shown based on 1000 bootstrap resamples, and 95% confidence limits are shown around the regression line. Green circle are New World sites, dark triangles are Old World in both panels.

presented ($F_{1,11} = 0.79$, $R^2 = 0.68$, $P < 0.001$). Dynamics for the 95th and 99th quantile are shown in Fig. S6. The former (95th) changes significantly with latitude ($F_{1,11} = 17.43$, $R^2 = 0.61$, $P = 0.0016$), while the 99th does not ($F_{1,11} = 0.79$, $R^2 = 0.067$, $P = 0.39$). The latter result (for the 99th quantile) is consistent with the static nature of the far reach of the tail of the distribution (see also the maximum observations in Fig. 2*B*, which do not change with latitude).

Appendix S4, Plant diversity and specialization

Multiple regression and path analysis were used to investigate relationships among plant diversity, diet breadth and relevant variables. First, we asked if there is a relationship between the species richness of plant families and the diet breadth of insects that attack those families. Table S4 shows results associated with particular factors in regression models, as described in the main text. Details for full models as follows: $F_{4,24} = 7.73$, $R^2 = 0.56$, $P < 0.001$ (for the model with all herbivores); $F_{4,21} = 2.01$, $R^2 = 0.28$, $P = 0.13$ (for the model with only herbivores occurring at sites $\leq$ 25 degrees of latitude); and $F_{4,21} = 3.67$, $R^2 = 0.41$, $P = 0.020$ (for herbivores occurring at sites >25 degrees of latitude). In these analyses, we included plant families from which caterpillars had been reared at least 100 times. Qualitatively similar results were obtained for all herbivores as well as for herbivores at lower latitudes (Table S4); much lower power was available for the higher latitude subset.

Finally, path analysis was used to address relationships among plant richness, latitude and dietary specialization for the thirteen Lepidoptera sites (dietary specialization was represented by α, the shape parameter from the Pareto distribution; higher
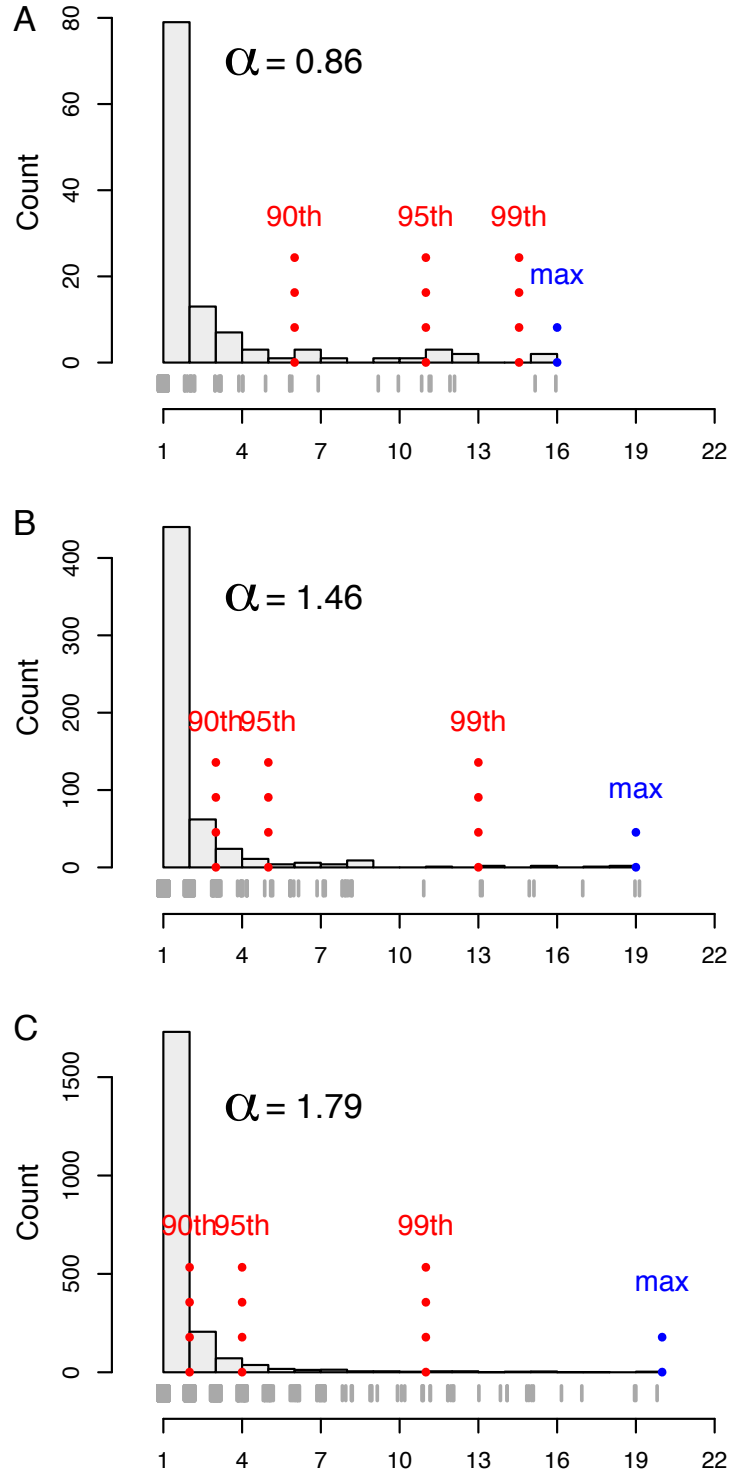
Fig. S5. Illustration of descriptive statistics associated with family–level diet breadth distributions from three sites: Connecticut (A), Brazil (B), and Ecuador (C).

625 values of α correspond to a greater fraction of specialized herbivores). Plant richness in these
626 analyses refers to the numbers of plant families and species encompassed by insect sampling at
627 each site. Results from analyses
628 are summarized in Fig. S7 for
629 two models, one in which plant
630 richness refers to plant species
631 ($\chi^2 = 1.62$, d.f. = 2, $P = 0.44$),
632 and another in which plant
633 richness refers to plant families
634 ($\chi^2 = 1.61$, d.f. = 2, $P = 0.45$).
635 Both models fit the data: P
636 values did not reject the null
637 hypothesis of fit. Multiplying
638 standardized path coefficients,
639 we see that the indirect effect
640 of latitude on specialization via
641 plant richness (in the model
642 with plant species richness) is -
643 0.70 * 0.22 = -0.15. Thus the
644 contribution of plant richness
645 to the latitudinal gradient in
646 specialization is approximately
647 1/4 the direct effect of latitude
648 on specialization (-0.65).
649 Analyses shown in Fig. S7 only
650 involved α calculated for
651 family-level diet breadth.
652 Analyses are not shown for
653 species-level α, for which
654 results were qualitatively
655 identical to the results in Fig.
656 S7.
657
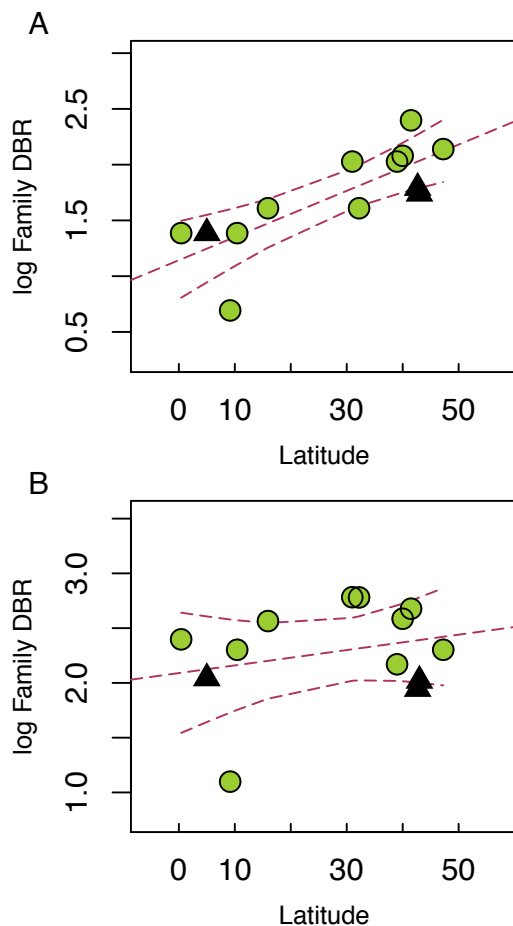
Fig. S6. 95th (A) and 99th (B) quantiles versus latitude; see main text and Fig. 2B for 90th quantile. 95% confidence limits are shown. As in other plots, green circles are New World, black triangles Old World.

Table S4. Results from three multiple regression models with independent variables as shown below, and the dependent variable of median species-level diet breadth for insects associated with different host plant families. The three analyses are for all insect herbivores, herbivores at and below 25 degrees of latitude, and herbivores at sites greater than 25 degrees.

Factor	Estimate	Std. error	T
All herbivores			
Richness	-0.36	0.067	-5.42 ***
Range	0.0076	0.0030	2.49 *
Age	0.49	0.30	1.61
Sample size	0.067	0.14	0.47
Tropical			
Richness	-0.14	0.059	-2.37 *
Range	0.0035	0.0027	1.30
Age	0.28	0.27	1.07
Sample size	-0.09	0.26	-0.74
Temperate			
Richness	-0.040	0.069	-0.59
Range	-0.0022	0.0032	-0.67
Age	0.17	0.32	0.54
Sample size	-0.37	0.15	-2.53*
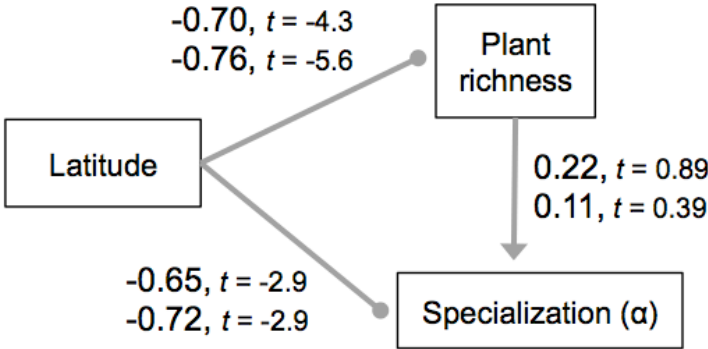
Fig. S7. Path analyses of latitude, plant richness and dietary specialization across all 13 focal sites; specialization is represented by α, the shape parameter from the Pareto distribution fit to family-level diet breadth. Next to each path, two sets of coefficients and test statistics are reported: the first value is for models that included plant species richness, and the second value is for models that included plant family richness. R^2 from both models as follows: 0.49 and 0.57 for plant richness in models with plant species richness and plant family richness, respectively; 0.68 and 0.66 for diet breadth, similarly for models with plant species and plant family richness.

727 *Appendix S5, R code for fit of the discrete, truncated Pareto*

728

729 ## To accompany Forister et al. "The global distribution of diet breadth in insect herbivores"
730 ## 11 June 2014
731 ## Two blocks of code below, one for estimation of Pareto fit,
732 ## and one for graphing a survival plot (as in Fig 1*D* and *E* of main text).

733

734 ## The Pareto fit works on a simple vector of diet breadth values (positive integers),
735 ## Usage:
736 ## dtparlogest(exampleData$sdbr,0,10)
737 ## where sdbr is the vector of diet breadth values, and 0 and 10 set bounds on
738 ## estimation of the shape parameter alpha.

739

740 ## The survival plotting function works on a data frame with two columns:
741 ## one with site identifier (collection locale, for example), and the second
742 ## with diet breadth values (each row being comprised of the diet breadth
743 ## for one herbivore species). Usage:
744 ## Pfit2("sdbr",exampleData,"siteName",-35,10,0,5)
745 ## where the first term specifies the diet breadth column,
746 ## the second term is the name of the data frame,
747 ## the third term is the name of the site (or rows within the data frame to be analyzed),
748 ## and the last terms set x and y limits for survival plot.

749

750 ################ main Pareto estimation ###############

751

752 ## Function computing least squares distance between logs of sample and
753 ## model cdf
754 lslogdist.dtpar<-function(x, a)
755 {
756   n <- length(x)
757   x <- sort(x)
758   xmin<-min(x)
759   xmax<-max(x)
760   xval<-seq(from=xmin, to=xmax, by=1)
761   nxval<-length(xval)
762   cdf<-rep(0, times=nxval)
763   for (i in 1:nxval){
764   cdf[i]<-length(x[x<=xval[i]])/n
765 }

766

767 g<-xmin
768 nu<-xmax
769 values=seq(from=g, to=nu, length=nu-g+1)
770 den=(1/g)^a - (1/(nu+1))^a
771 numvalues<-nu-g+1
772 numterm1=1/values^a

```r
773    numterm2<-rep(0, times=numvalues)
774    numterm2[1:numvalues-1]<-numterm1[2:numvalues]
775    numterm2[numvalues]<-1/(nu+1)^a
776    numterm2<-1/(values+1)^a
777    numterm2
778    num=numterm1-numterm2
779    probs<-num/den
780    sum(probs)
781    cprobs<-cumsum(probs)
782
783    ## survival function
784    surv<-1-cprobs
785    emp.surv<-1-cdf
786    lsurv<-length(surv)
787    lsurvm1<-lsurv-1
788    logsurv<-log(surv[1:lsurvm1])
789    logemp.surv<-log(emp.surv[1:lsurvm1])
790    lslogdist<-sum((logsurv-logemp.surv)^2)
791    lslogdist
792    }
793
794    dtparlogest<-function(y, int1, int2)
795    {
796            gamma <- min(y)
797            nu <- max(y)
798            out <- optimize(lslogdist.dtpar, interval = c(int1, int2), maximum = F,
799                    tol = 1e-005, x = y)
800            alpha <- out$minimum
801            cat("gamma=", gamma, "nu=", nu, "alpha=", alpha, "\n")
802            alpha
803    }
804
805    ################  survival plot ###############
806
807    Pfit2 <- function(type,data,where,ylim1,ylim2,xlim1,xlim2){
808            new<-subset(data,site==where)
809            x<-new[,type]
810            n <- length(x)
811            x <- sort(x)
812            xmin<-min(x)
813            xmax<-max(x)
814            xval<-sort(unique(x))
815            nxval<-length(xval)
816            nxvalm1<-nxval-1
817            surv<-rep(0, times=nxval)
818            for (i in 1:nxval){
```

```
819         surv[i]<-length(x[x>xval[i]])/n
820     }
821     a<-dtparlogest(x, 0.1, 10)
822     g<-xmin
823     nu<-xmax
824     values=seq(from=g, to=nu, length=nu-g+1)
825     xs<-1/(xval+1)^a
826     xs1<-xs -1/(nu+1)^a
827     num<-xs1[1:nxvalm1]
828     den<-1/g^a
829     den<-den-1/(nu+1)^a
830     surv.th<-num/den
831     xcoord<-log(xval[1:nxvalm1])
832     survth<-log(surv.th[1:nxvalm1])
833     survdata<-log(surv[1:nxvalm1])
834     plot(xcoord, survdata, col=c(1), pch=20,cex=1.5, xlim=c(xlim1, xlim2), ylim=c(ylim1,
835     ylim2),ylab="",xlab="")
836     points(log(xval[1:nxvalm1]), survth,pch=1,cex=2)
837     lpois<-mean(x)
838     cdf.pois<-ppois(xval, lpois)
839     surv.pois<-1-cdf.pois
840     log.surv.pois<-log(surv.pois[1:nxvalm1])
841     points(log(xval[1:nxvalm1]), log.surv.pois, col=1, pch=0,cex=1.5, xlab="", ylab="")
842     geo.mean<-1/mean(x)
843     xvalm1<-xval-1
844     cdf.geom<-pgeom(xvalm1, geo.mean)
845     surv.geom<-1-cdf.geom
846     log.surv.geom<-log(surv.geom[1:nxvalm1])
847     points(log(xval[1:nxvalm1]), log.surv.geom, cex=1.5,col=1, pch=2, xlab="", ylab="")
848     }
849
850
```

References

1. Wood D, *et al.* (2010) Tachinidae (tachinid flies, parasitic flies). *Manual of Central American Diptera* 2:1343-1417.
2. McGugan BM (1958) The Canadian forest insect survey. *Proceedings of the 10th International Congress of Entomology, Montreal*, pp 219-232.
3. McGugan BM (1958) *Forest Lepidoptera of Canada. Volume 1 Papilionidae to Arctiidae.* Forest Biology Division. Canadian Department of Agriculture 1034:1-76.
4. Prentice R (1962) *Forest Lepidoptera of Canada. Volume 2. Nycteolidae, Noctuidae, Notodontidae, and Liparidae.* Forest Biology Division. Canadian Department of Agriculture 1142:77-281.
5. Prentice R (1963) *Forest Lepidoptera of Canada. Volume 3. Lasiocampidae, Thyatiridae, Drepanidae, and Geometridae.* Forest Biology Division. Canadian Department of Agriculture 1013.
6. Prentice R (1965) *Forest Lepidoptera of Canada. Volume 4. Microlepidoptera.* Forest Biology Division. Canadian Department of Agriculture. 1142:544-840.
7. Wagner DL (2005) *Caterpillars of Eastern North America: A guide to identification and natural history* (Princeton Field Guides) (Princeton University Press, Singapore).
8. Wagner DL, Giles V, Reardon RC, & McManus ML (1998) Caterpillars of Eastern Forests. *USFS Technology Transfer Bulletin* FHTET-96-34.
9. Wagner DL, Ferguson DC, McCabe TL, Reardon RC (2002) Geometrid Caterpillars of Northeastern and Applachian Forests. *USFS Technology Transfer Bulletin* FHTET-2001-10.
10. Wagner DL, Schweitzer DF, Sullivan JB, Reardon RC (2011) *Owlet caterpillars of eastern North America* (Princeton University Press).
11. Novotny V, et al. (2006) Why are there so many species of herbivorous insects in tropical rainforests? *Science* 313(5790):1115-1118.
12. Covell CV, Jr. (2005) *A field guide to the moths of eastern North America* (Virginia Museum of Natural History, Martinsville, CA).
13. Holdridge LR, Grenke W (1971) *Forest environments in tropical life zones: a pilot study* (Pergamon Press, Oxford).
14. Croat TB (1978) *Flora of Barro Colorado Island* (Stanford University Press).
15. Leigh EG (1999) *Tropical Forest Ecology: A View from Barro Colorado Island* (Oxford University Press).
16. Novotny V, Basset Y, Miller SE, Drozd P, Cizek L (2002) Host specialization of leaf-chewing insects in a New Guinea rainforest. *J. Anim. Ecol.* 71(3):400-412.
17. Novotny V, et al. (2004) No tree an island: the plant-caterpillar food web of a secondary rain forest in New Guinea. *Ecol. Lett.* 7(11):1090-1100.
18. Hrcek J, Miller SE, Whitfield JB, Shima H, Novotny V (2013) Parasitism rate, parasitoid community composition and host specificity on exposed and semi-concealed caterpillars from a tropical rainforest. *Oecologia* 173(2):521-532.
19. Novotny V, et al. (2010) Guild-specific patterns of species richness and host specialization in plant-herbivore food webs from a tropical forest. *J. Anim. Ecol.* 79(6):1193-1203.
20. Laidlaw M, Kitching R, Goodall K, Small A, Stork N (2007) Temporal and spatial variation in an Australian tropical rainforest. *Austral Ecol.* 32(1):10-20.
21. Paijmans K (1976) *New Guinea vegetation* (Elsevier Scientific Publishing Company).
22. Whitfeld TJ, et al. (2012) Predicting tropical insect herbivore abundance from host plant traits

947　　　and phylogeny. *Ecology* 93(sp8):S211-S222.

948　23.　Novotny V, et al. (2012) Insects on plants: explaining the paradox of low diversity within
949　　　specialist herbivore guilds. *Am. Nat.* 179(3):351-362.

950　24.　Weiblen GD, Webb CO, Novotny V, Basset Y, Miller SE (2006) Phylogenetic dispersion of host
951　　　use in a tropical insect herbivore community. *Ecology* 87(sp7):S62-S75.

952　25.　Lewis OT, et al. (2002) Structure of a diverse tropical forest insect-parasitoid community. *J.*
953　　　*Anim. Ecol.* 71(5):855-873.

954　26.　Baje L, Stewart AJ, Novotny V (2014) Mesophyll cell-sucking herbivores (Cicadellidae:
955　　　Typhlocybinae) on rainforest trees in Papua New Guinea: local and regional diversity of a
956　　　taxonomically unexplored guild. *Ecol. Entomol.* 39(3):325-333.

957　27.　Dem FF, Stewart AJ, Gibson A, Weiblen GD, Novotny V (2013) Low host specificity in species-
958　　　rich assemblages of xylem-and phloem-feeding herbivores (Auchenorrhyncha) in a New Guinea
959　　　lowland rain forest. *J. Trop. Ecol.* 29(06):467-476.

960　28.　Michalska Z (1976) Owądy minujace Białowieskiego Parku Narodowego [Mining insects of the
961　　　National Park of Białowieża]. *Prace Komisji Matematyczno-Przyrodniczej. Seria B, Nauki*
962　　　*biologiczne* 44:1-81.

963　29.　Claridge M, Wilson M (1981) Host plant associations, diversity and species-area relationships of
964　　　mesophyll-feeding leafhoppers of trees and shrubs in Britain. *Ecol. Entomol.* 6(3):217-238.

965　30.　Nickel H (2003) *The leafhoppers and planthoppers of Germany (Hemiptera, Auchenorrhyncha):*
966　　　*patterns and strategies in a highly diverse group of phytophagous insects* (Pensoft, Sofia,
967　　　Moskau).

968　31.　Nickel H (2008) Tracking the elusive: leafhoppers and planthoppers in tree canopies of European
969　　　deciduous forests. *Canopy arthropod research in Europe: basic and applied studies from the*
970　　　*high frontier*, eds Floren A, Schmidl J (Bioform, Nürnberg), pp 175-214.

971