

Phylogenetic Analysis of Trophic Associations

A. R. Ives^{1,*} and H. C. J. Godfray^{2,†}

1. Department of Zoology, University of Wisconsin, Madison, Wisconsin 53706;

2. Natural Environment Research Council Centre for Population Biology, Division of Biology, Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, United Kingdom

Submitted July 29, 2005; Accepted February 27, 2006;

Electronically published May 19, 2006

Online enhancement: Matlab zip file.

ABSTRACT: Ecologists frequently collect data on the patterns of association between adjacent trophic levels in the form of binary or quantitative food webs. Here, we develop statistical methods to estimate the roles of consumer and resource phylogenies in explaining patterns of consumer-resource association. We use these methods to ask whether closely related consumer species are more likely to attack the same resource species and whether closely related resource species are more likely to be attacked by the same consumer species. We then show how to use estimates of phylogenetic signals to predict novel consumer-resource associations solely from the phylogenetic position of species for which no other (or only partial) data are available. Finally, we show how to combine phylogenetic information with information about species' ecological characteristics and life-history traits to estimate the effects of species traits on consumer-resource associations while accounting for phylogenies. We illustrate these techniques using a food web comprising species of parasitoids, leaf-mining moths, and their host plants.

Keywords: host-parasitoid interactions, community structure, predator-prey associations, phylogeny, comparative methods.

Why are some plant species attacked by many more species of herbivore than others (Strong et al. 1984)? And why do some parasite species have a much broader host range than more specialized species (Price 1980; Combes 2001)? These are examples of questions that can be asked of data sets that describe how two adjacent trophic levels in a food web are linked together. Not all food webs have species

organized in distinct trophic levels, but many do, or do to a good approximation, and increasingly quantitative data are becoming available that describe large sets of interacting species (van Veen et al. 2005).

Most analyses of these types of questions have either assumed that data from different species are statistically independent or considered the possibility of phylogenetic correlation at one trophic level but not the other. As is now widely appreciated, there is a danger of assuming that data from species are statistically independent; closely related species may share traits simply because they have inherited them from a common ancestor (Felsenstein 1985; Harvey and Pagel 1991). For example, swallows and martins (*Hirundinae*) reuse their nests year after year and are attacked by a relatively large number of species of insect ectoparasites (Rothschild and Clay 1952). However, this may have nothing to do with nesting habits and just be due to the common lousy ancestor of the clade; some other unmeasured trait might cause the high ectoparasite burden, with all swallows and martins sharing this unfortunate trait due to their common phylogenetic history. As Charles Darwin pointed out, the traits of species often reflect their evolutionary history, and when species traits affect the range of hosts they attack or the range of parasites that attack them, phylogenetic patterns of host-parasite associations are inevitable.

For analyses restricted to one trophic level, a wide range of statistical techniques is now available to estimate the extent of phylogenetic signal and to accommodate phylogenetic correlation in hypothesis testing. There are, however, few methods that simultaneously allow phylogenetic information from both trophic levels to be incorporated into the analysis of patterns of association. Although Page and colleagues (Page 1994, 2000; Charleston 1998; Page et al. 1998) developed methods to detect cocoladogenesis (cospeciation) in ectoparasites of birds and similar problems (Charleston 1998), here the parasites are monophagous or restricted to small clades of hosts, and phylogenetic processes dominate. Their techniques are not designed to explore patterns in more connected assemblages, where ecological determinants of trophic links may be as strong as or stronger than the effects of phylogeny.

The aim of this article is to develop methods for as-

* Corresponding author; e-mail: arives@wisc.edu.

† E-mail: c.godfray@imperial.ac.uk.

sessing the strength of phylogenetic signal that determines the patterns of association between consumers and resources in simple food webs. Specifically, we ask whether a given resource species is more likely to be used by phylogenetically related species of consumers and whether a given consumer species is more likely to use phylogenetically related resource species. Our objective is to identify the presence of phylogenetic signal (Blomberg et al. 2003) without attempting to ascribe any adaptive explanation; identifying phylogenetic correlations is a statistical issue, whereas ascribing adaptive explanations goes far beyond statistics (Blomberg and Garland 2002). We also ask whether it is possible to predict how a novel introduced species will fit into the food web, based on its phylogenetic position. Our motivation here is to help predict which native resource species will probably be at risk from an invasive consumer and which native consumers may be effective control agents for invasive resource species. Finally, we develop methods for regression analyses in which the strength of consumer-resource associations is regressed on life-history or environmental factors of consumer and/or resource species while incorporating phylogenetic signals. Incorporating phylogenetic signals into regression analyses will lead to the most precise estimates of regression coefficients. Throughout, we take a statistical approach, so parameter estimates and predictions are made with confidence intervals.

To illustrate and test the techniques for analyzing the cophylogenies of resources and consumers, we explore the patterns of association in a quantitative host-parasitoid food web. Parasitoids are insects that lay their eggs in or on the bodies of other insects, their hosts, which provide all the resources the developing parasitoid needs to reach maturity (Godfray 1994; Quicke 1997). Because of parasitoids' importance in biological control, there are abundant data on the parasitoid species complement of different hosts, and fewer but still numerous data on the number of host species attacked by different parasitoids. Analysis, largely without incorporating phylogenetic information, has suggested a variety of possible patterns, such as the effect of host geographical range on the number of parasitoid species that attack them and the effect of parasitoid feeding strategy (internally or externally on the host) on the host species they attack (Hawkins 1994). Recently, a number of quantitative host-parasitoid food webs have been constructed in which not only are the suitable hosts established for each parasitoid species but the strength of interaction between each host-parasitoid pair is quantified by the parasitism rates (e.g., Müller et al. 1999; Lewis et al. 2002). These quantitative food webs can be used to explore community structure and the potential for indirect interactions. Because these webs describe the densities of hosts, parasitoids, and their associations in the same units,

they are ideal for the quantitative exploration of the strength of phylogenetic signal in determining food web structure.

The food web we study includes 12 leaf-mining moths and 27 species of parasitoid wasps. We first ask whether there is phylogenetic signal in the observed associations between hosts and parasitoids and whether this signal acts through the host and/or parasitoid phylogeny. In addition to considering the phylogeny of the leaf miners, we also ask whether the host-parasitoid associations could be explained by the phylogeny of the plant species used by the leaf miners. This would occur if the parasitoid host range and foraging strategy are determined, at least in part, by the plant species used by the leaf miners (Vinson 1997). Once parameters for the phylogenetic signals are estimated, it is possible to use the phylogenies (discounted by the estimates of signal strength) to predict the host range for a novel parasitoid from the parasitoid's location on its phylogenetic tree and to predict the potential parasitoids of a novel host from the host's location on its phylogenetic tree. Finally, we consider factors that may influence the susceptibility of hosts to parasitism and factors that may influence the efficacy of parasitoids searching for hosts. Specifically, we ask whether hosts with broader geographical ranges experience parasitism from a greater number of parasitoid species (Hawkins and Lawton 1987) and whether parasitoids that feed externally on hosts (ectoparasitoids) attack a greater number of host species than those that feed internally on still-living hosts (endoparasitoids; Askew and Shaw 1986; Sheehan and Hawkins 1991).

Methods

Comparative Methods for Interacting Trophic Levels

Consider information on the strengths of association between n resource species and m consumer species in a food web or similar data structure. For narrative simplicity, we henceforth call the resources "hosts" and the consumers "parasitoids." Let the association strengths (whose exact form we need not specify for the moment) be denoted A_{ik} ($i = 1, \dots, n$; $k = 1, \dots, m$). The total number of associations (including zeros) is $N = nm$.

Assume that there is a host trait X_h and a parasitoid trait X_p that together affect the hosts' susceptibility to parasitism and the parasitoids' effectiveness at finding and attacking hosts. These traits are assumed to depend on phylogenetic relatedness so that closely related host or parasitoid species are more likely to have similar values of X_h or X_p , respectively. We assume that phylogenetic relatedness gives rise to covariance matrices \mathbf{V} and \mathbf{U} , whose elements v_{ij} and u_{kl} describe the covariance in the trait

values of X_h and X_p for pairs of host and parasitoid species; note that for this indexing system, sequential letters are used for hosts (i and j) and parasitoids (k and l). For our analyses, we standardize the variances in the covariance matrices \mathbf{V} and \mathbf{U} by setting all diagonal elements equal to the maximum diagonal element and then dividing all elements in the matrix by this maximum; this is equivalent to extending the tip branch lengths for species to make them all contemporaneous in the phylogenetic tree. Contemporaneous tips imply that the variance in trait values is the same for all species in the phylogenetic tree (Grafen 1989; Harvey and Pagel 1991), thereby making \mathbf{V} and \mathbf{U} correlation matrices. This standardization, however, is not needed for any of the statistical computations; the techniques we present can be applied to any type of covariance matrices \mathbf{V} and \mathbf{U} . Therefore, if there is good statistical or biological reason to weight species differently or use phylogenetic trees with noncontemporaneous tips, the procedures can be implemented without modification.

We assume that the strength of the host-parasitoid association is given by the product of the values of the host and parasitoid traits X_h and X_p , so the strength of association between the i - k and j - l host-parasitoid pairs is $v_{ij}u_{kl}$. Therefore, for a given host species i , the phylogenetic correlation between attacks from parasitoids k and l is given by the parasitoid phylogeny u_{kl} (since $v_{ii} = 1$). Similarly, for a given parasitoid species k , the phylogenetic correlation between attacks on hosts i and j is given by the host phylogeny v_{ij} . If both hosts and parasitoids are phylogenetically related, then attacks between the i - k and j - l host-parasitoid pairs are correlated. However, a positive covariance between the strengths of association of two host-parasitoid pairs occurs only if both the host and parasitoid species are phylogenetically related ($v_{ij} > 0$ and $u_{kl} > 0$). This produces a joint evolutionary model because traits X_h and X_p must both evolve down the host and parasitoid phylogenies for strengths of association between different host-parasitoid pairs to be correlated. Additional models for combining host and parasitoid phylogenies are possible, and the statistical techniques presented below can be deployed for any reasonable model.

The strength of phylogenetic association between the n host and m parasitoid species in the web can be modeled statistically as follows. First, take the $n \times m$ matrix with elements A_{ik} and rearrange it into the $N \times 1$ vector \mathbf{A} by “stacking” columns; $\mathbf{A} = [A_{1,1}, A_{2,1}, \dots, A_{n,1}, \dots, A_{n,m}]'$, where the prime denotes transpose. Let

$$\mathbf{A} = \mathbf{b}_0 + \boldsymbol{\varepsilon}, \quad (1)$$

where \mathbf{b}_0 is the phylogenetically corrected mean of association strength, and $\boldsymbol{\varepsilon}$ is an $N \times 1$ vector of zero-mean random variables (not necessarily normally distributed)

with covariance matrix $E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \mathbf{W}$. The phylogenies of hosts and parasitoids are incorporated into the matrix \mathbf{W} . Specifically, for the evolutionary model described above, $\mathbf{W} = \mathbf{U} \otimes \mathbf{V}$, where \otimes is the Kronecker (outer) product.

Measuring phylogenetic signal. The strength of phylogenetic signal depends on the magnitudes of the covariances in the host and parasitoid covariance matrices \mathbf{V} and \mathbf{U} , with greater values of v_{ij} and u_{kl} corresponding to stronger phylogenetic signal through the host and parasitoid phylogenies, respectively. A common assumption in phylogenetic analyses is that the evolution of a trait on a phylogenetic tree can be described as a Brownian motion process in which the value of a trait changes randomly and incrementally through time. In this case, the distribution of trait values among species follows a multivariate normal distribution, with the covariance in trait values between two species being proportional to the branch length of their shared phylogeny (Felsenstein 1985; Martins and Hansen 1997; Garland and Ives 2000). Here we use a more flexible model based on an Ornstein-Uhlenbeck (OU) process that can incorporate stabilizing selection. Specifically, Blomberg et al. (2003) derived a version of the OU process in which there is a single optimum value of a trait for all species in the phylogeny. The stronger the selection toward the global optimum, the weaker the phylogenetic signal because phylogenetic correlations are degraded through time as species traits are dictated more by selection. This version of the OU process leads to a formula in which a parameter d determines the strength of phylogenetic signal, with $d = 1$ (no selection) corresponding to the Brownian motion assumption, $0 < d < 1$ representing stabilizing selection, $d = 0$ depicting the absence of phylogenetic correlation (a “star” phylogeny), and $d > 1$ corresponding to disruptive selection, in which the phylogenetic signal is amplified (app. A). Although we use this formula to estimate phylogenetic signal, we do not mean to ascribe any observed phylogenetic signal to stabilizing selection; instead, we use the OU formulation to give a reasonable statistical description of phylogenetic signal that could be produced by any number of processes.

To allow for differing strengths of signal in the host and parasitoid phylogenies, we estimate d_h and d_p as separate parameters for host and parasitoid lineages to give the covariance matrices $\mathbf{V}(d_h)$ and $\mathbf{U}(d_p)$. While several estimation procedures are possible (e.g., maximum likelihood, restricted maximum likelihood), here we use a version of estimated generalized least squares that does not assume the distribution of association strength among hosts and parasitoids is normally distributed (Judge et al. 1985). We then used bootstrapping (Efron and Tibshirani 1993) to obtain confidence intervals of the parameter estimates (app. A).

Predicting the strength of novel associations. Suppose that

the phylogenetic position of a host is known, but no information is available about the strengths of association with parasitoids in the community. In appendix B we show how these strengths can be estimated (with confidence limits), and similar formulas can be derived for novel parasitoids. Furthermore, it is possible to make more informed predictions. Suppose, for example, for a new host species, one knows the parasitoid attack rate by some but not all parasitoid species. The known information can be incorporated into the estimation procedure, and the strength of unknown associations can be estimated.

Analyses with covariates. The statistical model for host-parasitoid associations in equation (1) depends simply on the phylogenetic signal. However, information about phylogenetically correlated or uncorrelated traits that may help explain the strengths of association can also be incorporated in the model, and their explanatory power may be tested statistically. This leads to a regression analysis in which host-parasitoid association strength is regressed on independent variables, with phylogenetic signal included in the process error (residual) terms (e.g., Felsenstein 1985, 2004; Grafen 1989; Harvey and Pagel 1991; Garland et al. 1993; Martins and Hansen 1997). The statistical motivation for incorporating phylogenetic signal is that this produces the “best” (minimum variance) estimates of the regression coefficients (Judge et al. 1985; Rohlf 2001).

Suppose information is available about M measured traits. Let

$$\mathbf{A} = \mathbf{X}\mathbf{b} + \varepsilon, \quad (2)$$

where \mathbf{X} is a $N \times (M + 1)$ matrix whose first column contains 1's and whose remaining M columns contain values of covariates (independent variables). The vector $\mathbf{b} = [b_0, b_1, \dots, b_M]'$ contains regression coefficients, with b_0 denoting the intercept and b_q ($q = 1, \dots, M$) denoting the coefficients for the M covariates. Parameter estimation and calculation of confidence intervals proceed in the same way as for equation (1) (app. A).

Tests with a Leaf Miner–Parasitoid Web

We explore the use of the comparative methods developed above using data from a food web that describes the associations between 12 leaf-mining moths in the genus *Phyllonorycter* (Gracillariidae) and the 27 species of parasitoid wasps (Hymenoptera) that attack them. Full details of the construction of the web and the species involved are given in Rott and Godfray (2000). Briefly, the study site was a 10,000-m² area of woodland containing four species of trees. The trees were sampled twice each in 1992 and 1993 at times corresponding to the two annual generations of the leaf miners. A total of 8,242 miners were

collected and reared in the laboratory, and the parasitoids were identified. The leaf miners collectively used all four tree species, *Alnus glutinosa* L., *Salix cinerea* L., *Quercus robur* L., and *Betula pubescens* L., but each leaf miner species was found on only a single tree species.

Phylogenetic data. Phylogenies were available for all three trophic levels in the food web. The four species of host plant belong to four different genera whose phylogenetic relationships are well supported by both morphological and molecular data (Soltis et al. 2000). The 12 host species are all included in a phylogeny of the genus *Phyllonorycter* constructed by Lopez-Vaamonde et al. (2003) and based on molecular (28S rDNA) data. The parasitoids include 19 members of the family Eulophidae that were placed on a phylogeny largely using Gautier et al.'s (2000) study based on sequence data (again 28S rDNA) from members of all the genera found in the web. The other parasitoids were incorporated based on long-standing morphological hypotheses about hymenopteran phylogeny (Gauld and Bolton 1988). We recognize that the topologies of the host and parasitoid phylogenetic trees are not known with complete certainty and that if the uncertainty were large, it might affect results of phylogenetic analyses (Purvis and Garland 1993; Garland and Díaz-Uriarte 1999; Housworth and Martins 2001; Huelsenbeck and Rannala 2003). Nonetheless, we have high confidence in the topologies of our phylogenetic trees, and our results are unlikely to be affected by topological uncertainties.

Rott and Godfray (2000) noted that hosts attacking the same tree species often tended to have similar parasitoid complexes even if the hosts were not closely related. To test for a signal associated with the phylogeny of the tree species, we created a “pseudophylogeny” for the hosts based on the tree phylogeny. Specifically, we constructed the plant phylogeny with branch lengths selected to produce contemporaneous tips and then attached each host species to these tips to give a polytomy with branches equal to the shortest branch on the tree phylogeny.

A phylogenetic signal from the plant phylogeny on host-parasitoid associations could be a spurious result: if host-parasitoid associations have a phylogenetic signal from the host species, and if plant-host associations have a phylogenetic signal from the plant species, then one would expect host-parasitoid associations to be apparently related to the plant phylogeny. To guard against this, we compared the explanatory power of host and plant phylogenies. We assumed that the phylogenetic correlation between host species in trait X_h was given by the covariance matrix $\mathbf{V}(w) = \mathbf{V}_t + w\mathbf{V}_h$, where \mathbf{V}_h and \mathbf{V}_t are the covariance matrices derived from host and plant (pseudohost) phylogenies, assuming Brownian motion evolution, and w is a parameter that weights the importance of the two phylogenetic signals. We used this base matrix to construct $\mathbf{V}(d_h)$,

w) as done previously under the assumption of an OU process acting on the aggregate phylogeny given by $V(w)$. The relative explanatory power of the two phylogenies can then be tested by estimating w with its 95% confidence interval (app. A).

This procedure for separating the effects of host phylogeny from plant phylogeny is simple for our data set because each host species is only found on a single plant. Nonetheless, this approach can be modified for more general tritrophic systems as follows. Suppose each host was found on multiple plants. Thus, a given host i attacked by parasitoid k occurred on plants q and r . The association between host i and parasitoid k can be partitioned into the association strength of plant q , A_{ikq} , and the association strength of plant r , A_{ikr} , with $A_{ikq} + A_{ikr} = A_{ik}$. This partitioning of associations can be incorporated into the plant phylogenetic tree by introducing polytomies to each tip corresponding to a different host that is found on the plant; thus, the number of tips on the tree would equal the number of plant-host associations. Similarly, the host phylogenetic tree could be modified by introducing polytomies to each tip corresponding to a different plant used by the host. The resulting covariance matrices V_t and V_h could then be used as described above.

Strength of association. We used three different measures of the strength of the association between host and parasitoid species: (1) the overall parasitoid attack rate on each host species, a measure of the total mortality inflicted by a parasitoid influenced by both parasitoid abundance and selectivity; (2) the per capita parasitoid attack rate, simply the first measure divided by the abundance of a parasitoid species and hence a measure of selectivity; and (3) a binary measure of whether the host-parasitoid association had been observed. Which measure is most appropriate depends on the precise hypothesis that is being tested.

To derive the first measure, let H_i denote the number of hosts of species i (parasitized or not), and let F_{ik} denote the number of hosts of species i parasitized by species k . Assuming that parasitoids search for hosts of a given species randomly, the proportion of hosts escaping parasitism is the zero term of a Poisson distribution, and therefore

$$F_{ik} = H_i[1 - \exp(-A_{ik})], \quad (3)$$

where A_{ik} is the net attack rate of parasitoid k on host i . Rearranging equation (3) gives

$$A_{ik} = -\log\left(1 - \frac{F_{ik}}{H_i}\right). \quad (4)$$

Thus, A_{ik} is our first measure of association between hosts and parasitoids.

Because A_{ik} depends on both parasitoid selectivity and parasitoid abundance, to obtain the second measure of association a_{ik} that depends only on selectivity, we divided A_{ik} by the relative abundance of parasitoid species. We considered two methods of determining parasitoid abundance. The first is to use the average value of A_{ik} from all host species i from which parasitoid k was reared. In this case, the zeros for host species that the parasitoid does not attack are not included in the average. This assumes that the number of host individuals parasitized by a species is a good indicator of the population size of the parasitoid species. The second way to determine parasitoid abundance is to use the average value of A_{ik} calculated for all host species i , thereby including the zeros for host species that are not attacked by parasitoid k . Calculating the average value of A_{ik} in this way would produce higher estimates of the per capita attack rate for parasitoid species that attack fewer host species because factoring zeros into the estimates of the average value of A_{ik} and then dividing each A_{ik} by this average will lead to higher values of a_{ik} . These high values of a_{ik} would in effect assume that individuals from parasitoid species attacking fewer host species are “lost” when searching for unsuitable hosts. Due to this unrealistic assumption when averaging A_{ik} over all host species, we prefer the former procedure to obtain relative parasitoid abundances and calculate a_{ik} , and we present the results for this procedure here. Results for a_{ik} calculated by the alternative procedure are given in appendix C. Preliminary diagnostics (not presented) supported applying a square root transform to both A_{ik} and a_{ik} at the onset of the analyses.

For the third measure of association, we simply assigned a value of one to host-parasitoid pairs for which the parasitoid was reared from the host and zero otherwise. Using a binary dependent variable raises some technical problems with bootstrapping to obtain confidence intervals for parameter estimates, and these are discussed in appendix A.

Covariates. We investigated the influence of two covariates on the association between host and parasitoid species. We specifically wanted to test whether the two covariates affected whether parasitoid species attacked host species, and we consequently used the simple binary measure of association. The two covariates we investigated were the geographic range of a host species and whether the parasitoid was ecto- or endoparasitic. We anticipated a positive correlation between host geographical range and the number of parasitoids attacking the host. Broader geographical ranges expose host species to a greater diversity of parasitoids, and hence, hosts with broad ranges will probably have a greater number of parasitoids that have

evolved to attack them (Godfray 1994; Hawkins 1994). Host geographical range was measured by the number of Watsonian vice-counties (roughly equal units of area within Great Britain and Ireland) in which the host species was found in the British Isles by using range maps in Emmet et al. (1985). We anticipated that ectoparasitoids would have a broader host species range than endoparasitoids because they do not need to live within their still-living hosts and therefore require fewer adaptations to their host's physiology (Askew and Shaw 1986; Sheehan and Hawkins 1991; Godfray 1994).

Computer code written in Matlab (ver. 7.0, MathWorks 2005) for the core analyses in the article (table 1) is given in appendix D.

Results

Phylogenetic Signal

For both the parasitoid attack rate A_{ik} and the per capita parasitoid attack rate a_{ik} , we estimated the strengths of phylogenetic signal in the host and parasitoid phylogenies, d_h and d_p . For the host, we used both the host phylogeny to give the covariance matrix for host phylogenetic relatedness, V_h , and the phylogeny created by grafting host species onto the phylogeny of the plant species they use, V_t .

Both parasitoid attack rates A_{ik} and per capita parasitoid attack rates a_{ik} showed moderately strong phylogenetic signals through the host phylogeny, whereas the signal through the parasitoid phylogeny was weak or nonexistent (table 1). The overall strength of the phylogenetic signal can be assessed by comparing the mean squared error (MSE) calculated for the full model (MSE_d), the MSE derived under the assumption of no phylogenetic covariances (a "star phylogeny"; MSE_{star}), and the MSE derived assuming Brownian motion evolution (MSE_b ; app. A). Lower

values of MSE indicate better fit of the specific model to the data, in that the model with the lowest MSE leaves the smallest unexplained variance (app. A). The MSE can also be interpreted heuristically as an estimate of the rate of evolution (Blomberg et al. 2003); the lower the MSE, the less evolutionary change is needed to explain the observed variance in the data.

The values of MSE_d derived using the host plant covariance matrix V_t are smaller than those obtained from the host covariance matrix V_h (table 1). This suggests that there is a stronger signal from the plant phylogeny than the host phylogeny in the strength of host-parasitoid association. In assessing the relative strengths of the two, values of d_h cannot be compared directly. However, comparing the covariance matrices $V_h(d_h)$ and $V_t(d_h)$ at their respective optimal values of d_h indicates the strength of covariances estimated from the data. Figure 1 gives a visual depiction of the strength of signal in the value of A_{ik} from the host (fig. 1A) and plant (fig. 1B) phylogenies; the stronger signal from the latter is visible as longer shared branch lengths (higher covariances).

We derived a statistical test to compare the signals from the host phylogeny and the plant phylogeny using the covariance matrix $V = V_t + wV_h$ and estimating the value of w along with d_h (table 2). For A_{ik} , the estimate of w was 0, and for a_{ik} , the 95% confidence interval of w contained 0. These results show no statistical evidence for an effect of host phylogeny on the strength of host-parasitoid association after accounting for an effect of plant phylogeny.

The presence of a signal from the host phylogeny when plant phylogeny is not included (table 1) could be the result of phylogenetic correlation in association between host and plant species. We analyzed the pattern of plant-host association using the same approach as we used to analyze the strength of host-parasitoid association, al-

Table 1: Phylogenetic signal in parasitoid-host associations in which d_h and d_p measure the strength of signal from the host and parasitoid phylogenies, respectively

Dependent variable, host phylogeny	d_h	d_p	MSE_d	MSE_{star}	MSE_b
$\sqrt{A_{ik}}$:					
V_h	.40 (.09, .77)	0 (0, .14)	.010	.011	.012
V_t	1.30 (.90, 1.78)	.19 (0, .46)	.0052	.011	.0059
$\sqrt{a_{ik}}$:					
V_h	.45 (.25, .67)	.086 (0, .24)	.24	.27	.34
V_t	.63 (.45, .79)	.19 (.04, .35)	.20	.27	.25

Note: Two phylogenies were used for the hosts: the true host phylogeny V_h and the pseudophylogeny constructed from trees they attack, V_t . Dependent variables are the parasitoid attack rate A_{ik} and the per capita parasitoid attack rate a_{ik} , both square root transformed. Mean squared errors (MSE) are given for the cases when d_h and d_p have estimated values (MSE_d), when $d_h = 0$ and $d_p = 0$ (MSE_{star}), and when $d_h = 1$ and $d_p = 1$ (MSE_b). Ranges in parentheses are 95% approximate confidence bounds obtained from bootstrapping.

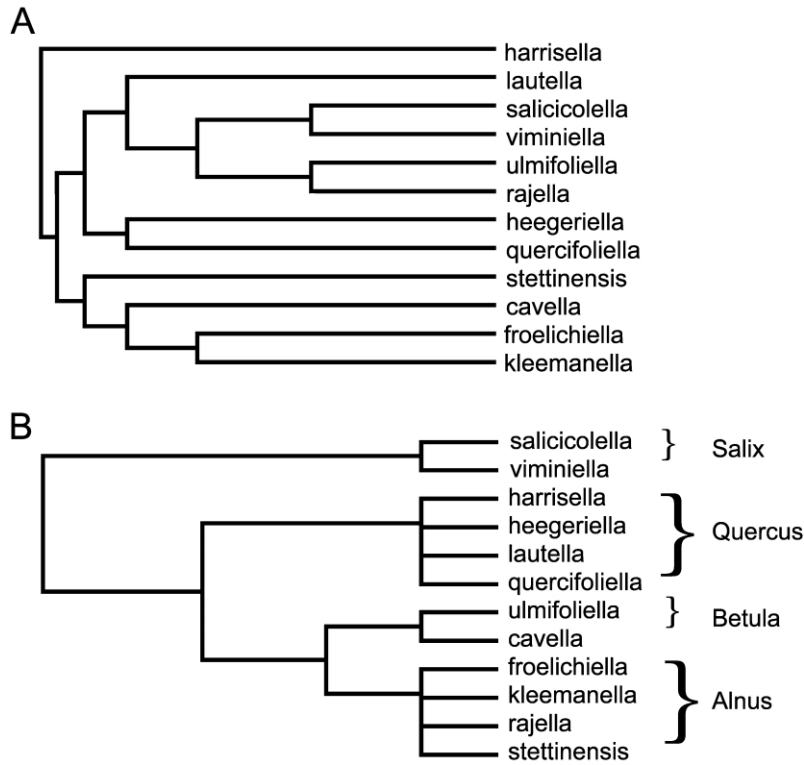


Figure 1: A, Phylogeny for host species of the genus *Phyllonorycter*, and B, phylogeny created by grafting host species onto the tips of the phylogeny of the tree species they attack. Branch lengths of the phylogenies were computed from the covariance matrices $V_h(d_h)$ and $V_t(d_h)$, with values of d_h fitted from the model of parasitoid attack rates, A_{ik} (table 1). The longer shared branch lengths of second phylogeny (B) indicate greater covariances and hence stronger phylogenetic signal from the phylogeny of host species grafted onto the phylogeny of trees they attack.

though we used a simple binary measure of whether the leaf miner fed on the tree species. There was no signal through the plant phylogeny ($d_t = 0$) and a weak signal through the host phylogeny ($d_h = 0.12$), although it was not statistically significantly different from 0. Therefore, there is no statistical evidence that the signal from the host phylogeny in the strength of host-parasitoid associations is driven solely by a phylogenetic association between hosts and plants. The relatively small sample size (four tree species), however, limits the statistical power of this test.

Predicted Associations for Novel Hosts or Parasitoids

When phylogenetic signals are strong, phylogeny can be used to predict the association between new hosts or parasitoids that are introduced into a community. To illustrate these predictions, consider the case of predicting the attack rate by all parasitoids in a community on a novel host species z , given by the vector A_z . Because the signal of the plant phylogeny is stronger than the signal of the host phylogeny, we used the covariance matrix V_t to predict associations for the novel host. To assess the accuracy of

the predictions, we let each host species in turn play the role of the novel species by removing it from the data set, fitting the values of d_h and d_p to the reduced data set, and then computing the estimates of association \hat{A}_z (eq. B1) using only information about the reduced assemblage of host species. We then computed the correlation between the estimates \hat{A}_z and the observed attack rates A_z . Repeating this by allowing each host in turn to be novel produced 12 sets of predictions and correlations. We then repeated this procedure for novel parasitoids, removing each of the 27 species, computing predicted per capita attack rates \hat{a}_z , and correlating these to the observed per capita attack rates a_z . Note that we predicted A_{ik} for novel hosts to give the overall effect of parasitism, and we predicted a_{ik} for novel parasitoids to give their per capita attack rates.

In figure 2A, 2B we plot A_z against \hat{A}_z for the two host species with the highest (0.96; *Phyllonorycter salicicolella*) and lowest (0.49; *Phyllonorycter kleemanella*) correlation coefficients. For all 12 species, the median value of the correlation coefficient was 0.91 (fig. 2C). Thus, predictions of the parasitoid attack rates on novel host species based on the phylogenetic relationship of their host plants are

Table 2: Phylogenetic signal from the tree pseudophylogeny versus the host phylogeny using the host covariance matrix $\mathbf{V} = \mathbf{V}_t + w\mathbf{V}_h$

Dependent variable	d_h	d_p	w	MSE _d	MSE _{star}	MSE _b
$\sqrt{A_{ik}}$	1.14 (.96, 1.31)	.19 (0, .45)	0	.0052	.011	.0078
$\sqrt{a_{ik}}$.87 (.71, .97)	.19 (.02, .33)	.092 (0, .33)	.20	.27	.27

Note: Dependent variables are the parasitoid attack rate A_{ik} and the per capita parasitoid attack rate a_{ik} , both square root transformed. Mean squared errors are given for the cases when d_h and d_p have estimated values (MSE_d), when $d_h = 0$ and $d_p = 0$ (MSE_{star}), and when $d_h = 1$ and $d_p = 1$ (MSE_b). Ranges in parentheses are 95% approximate confidence bounds.

quite good, as is expected from the strong phylogenetic signal of host plants.

We used the same procedure to predict the per capita attack rates of a new parasitoid species z on the existing host species in the community, $\hat{\mathbf{a}}_z$. The correlations between \mathbf{a}_z and $\hat{\mathbf{a}}_z$ ranged from -0.71 to 0.90 , with a median of 0.05 (fig. 2D–2F). These results reflect the signal of parasitoid phylogeny on the per capita parasitoid attack rates a_{ik} being weaker than that found for the plant phylogeny on the parasitoid attack rates A_{ik} (table 1).

Analyses with Covariates

Using the regression equation (2), we analyzed the effects of host geographical range and parasitoid feeding mode (endo- vs. ectoparasitism) on the pattern of association between hosts and parasitoids, where the pattern of association was coded by the binary 0 or 1. Contrary to expectation, the analyses failed to reject the null hypothesis that host geographical range had no effect on number of parasitoid species attacking a host species (table 3). Similarly, contrary to expectation, the analyses failed to reject the null hypothesis that ectoparasitoids had a wider host range than endoparasitoids (table 3). However, as anticipated by prior analyses (table 1), there was a strong signal through the plant phylogeny and a weaker signal through the parasitoid phylogeny on the binary measure of association. This is demonstrated by statistically significant positive values of d_h and d_p and values of MSE considerably lower than those obtained using a star phylogeny with 0 phylogenetic covariance between species.

Discussion

Patterns of association between adjacent trophic levels are one of the commonest types of data collected in ecology. Biologists and natural historians have amassed extensive data on the species of herbivores that attack different species of plants, the parasitoids that attack host insects, the parasites and pathogens found on different hosts, and many other types of association. In some cases, the consumer is highly specific, found attacking a single resource species and often with a phylogenetic history that is con-

gruent or nearly congruent with that of the resource (Combes 2001). However, much more commonly, the consumer is polyphagous, and there is poorer matching between the structure of the consumer and resource phylogenies (Price 1980). Previous analyses of patterns of consumer-resource association have either ignored phylogenetic correlations or considered just one trophic level, and they cannot accommodate information about the strength of association between consumers and resources beyond binary descriptions of whether an association exists. The aim of the work described here is to provide statistical methods to assess the strength of any phylogenetic signal and to account for it in the analysis of consumer-resource associations.

Our approach is based on established statistics of phylogenetic comparative methods (Felsenstein 1985; Garland et al. 1993; Martins and Hansen 1997; Garland and Ives 2000; Felsenstein 2004). The methods are relatively easy to apply using phylogenetic data that are becoming ever more readily available. The basic approach is to structure the problem of consumer-resource associations as a statistical model (eqq. [1], [2]) in which phylogenies are used to give the covariance structure of the “error” terms. Structured in this way, standard statistical formulas can be used to predict new values (association strengths between novel species) and estimate the effects of additional independent variables (ecological or life-history traits). Thus, the approach is flexible and can be modified to perform the range of tasks often demanded of regression analyses.

We emphasize that the statistical toolbox we have derived is intended to identify correlations in data sets that are consistent with patterns expected from phylogenies, but these should not be interpreted as tests of specific hypotheses of evolutionary processes (Martins 2000; Orzack and Sober 2001; Hansen and Orzack 2005). We have used the term “phylogenetic signal” to refer to a description of phylogenetic pattern divorced from any specific mechanism (Blomberg et al. 2003). If a phylogenetic signal is identified, however, this can be used as a license to hunt for possible adaptive (or nonadaptive) explanations of the patterns. In regression analyses with covariates (independent variables), incorporating phylogenetic signals will lead to better estimates of the regression coefficients, and

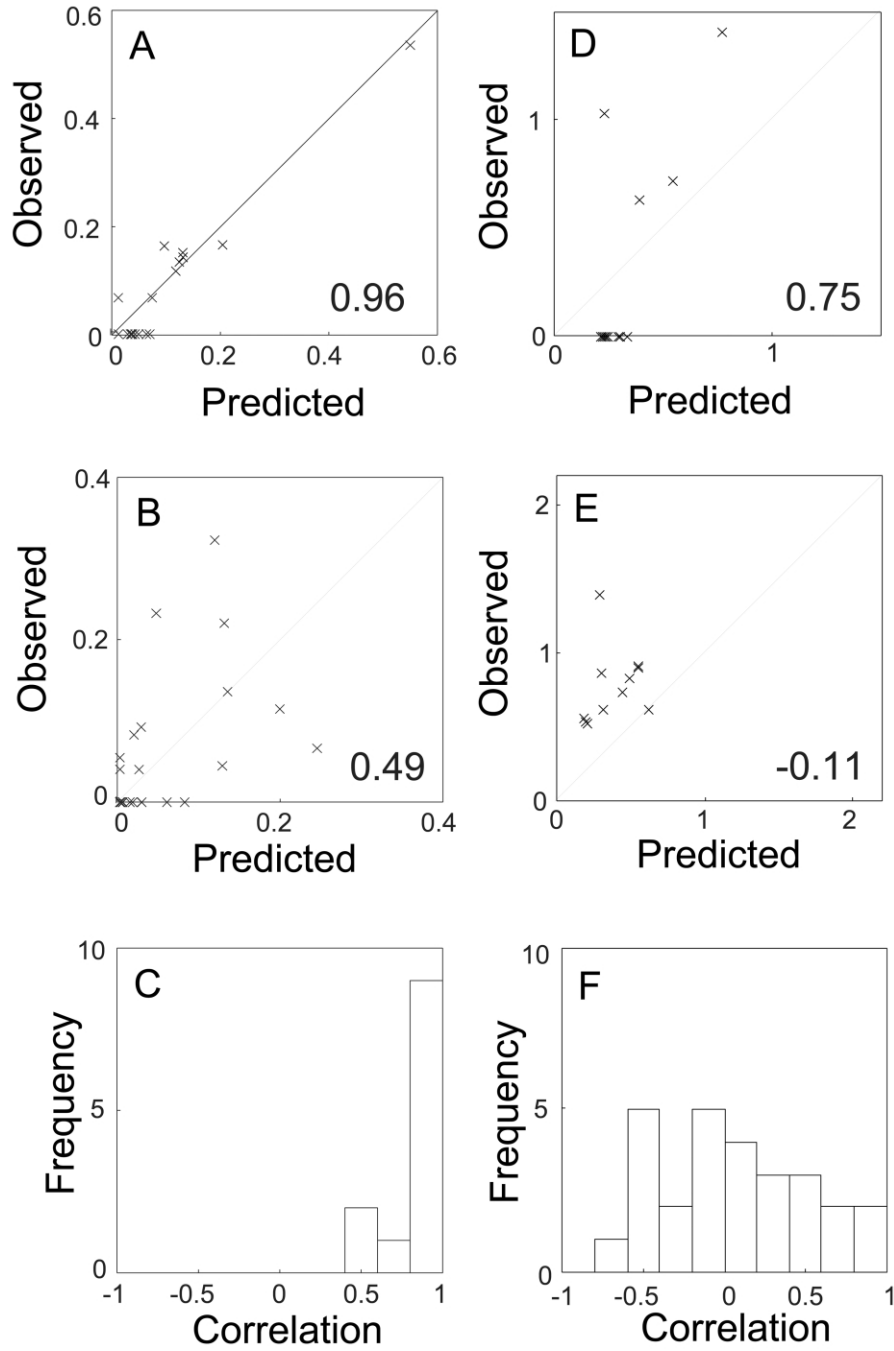


Figure 2: Observed versus predicted values of the parasitoid attack rates $(A_{ik})^{1/2}$ on the host species *Phyllonorycter viminiella* (A) and *Phyllonorycter stettinensis* (B). Predictions were made assuming that the only thing known about the host species was the phylogenetic relatedness of their host plants and the covariance matrix \mathbf{V}_r . Correlations between predicted and observed values are listed in each panel. C, Histogram of correlations between predicted and observed values when each of the 12 host species is treated as a new species; the median correlation is 0.91. Predicted versus observed values of the per capita parasitoid attack rates $(a_{ik})^{1/2}$ for *Pnigalio pectinicornis* (D) and *Achrysocharoides splendens* (E). F, Histogram of correlations between predicted and observed values when each of the 27 parasitoid species is treated as a new species; median correlation = 0.05.

Table 3: Phylogenetic regression of host-parasitoid association on either host geographic range or parasitoid ecto- versus endoparasitism

Independent variable, phylogenies	b_1	d_h	d_p	MSE
Host geographic range:				
V_p, U	1.04 (−.71, 2.68)	.923 (.667, 1.18)	.253 (.080, .403)	.152
Stars	.602 (−1.19, 2.39)	0	0	.242
Parasitoid ecto-/endoparasitism:				
V_p, U	.015 (−.189, .207)	.764 (.543, .950)	.258 (.078, .437)	.153
Stars	−.019 (−.126, .088)	0	0	.242

Note: The dependent variable is whether or not (1 or 0) a parasitoid species attacks a host species, and b_1 is the regression coefficient. Hosts were assumed to have the pseudophylogeny given by the trees they attack, V_p . Either the phylogenetic signal was estimated or star phylogenies were assumed by setting $d_h = 0$ and $d_p = 0$. Ranges in parentheses are 95% approximate confidence bounds.

therefore on strictly statistical grounds, phylogenetic signals should be investigated in studies of species associations (Ives and Zhu 2005). Detecting phylogenetic signals will depend on the sample size (number of possible consumer-resource associations); for example, for analyses of phylogenetic signals in single clades, Blomberg et al. (2003) found that at least 20 species were often needed to detect a phylogenetic signal with convincing statistical power. Our data set included 324 ($= 12 \times 27$) associations and had high statistical power. It is difficult to say how much smaller sample sizes could be while maintaining high statistical power because power will also depend on the strength of the true phylogenetic signal and the topology of phylogenetic trees; we suspect, however, that a minimum of six species in each clade is necessary.

For the specific example of leaf miners and their parasitoids that we analyzed, patterns of association showed a signal from the host phylogeny but only a weak or non-existent signal from the parasitoid phylogeny. Furthermore, the signal through the phylogeny of tree species attacked by the leaf miners was stronger than that of the host phylogeny itself. This indicates that leaf miner-parasitoid associations are strongly influenced by plants. Moreover, this effect is not just due to the tendency for phylogenetically related leaf miner species to feed on the same host plant because when statistically testing the simultaneous effects of host versus plant phylogenies, no statistically significant effect was found for the host phylogeny (table 2). There are several possible explanations for the strong effect of tree phylogeny. First, parasitoids might find leaf miners using cues that are associated with the tree species rather than the hosts per se, as is well-known for other host-parasitoid systems (Vinson 1997). Second, certain tree species might make hosts particularly vulnerable or resistant to specific foraging strategies employed by specific parasitoids. Finally, associations between leaf miners and trees could have created the evolutionary setting for associations between leaf miners and parasitoids to develop. For example, a parasitoid that initially specialized on a single leaf miner species might have evolved

host-range expansion to another leaf miner species that uses the same habitat (tree species); selection on parasitoids for host-range expansion should be greatest for hosts that are frequently encountered by the parasitoids (Quicke 1997).

The lack of signal through the parasitoid phylogeny surprised us; closely related parasitoid species were at most slightly more likely to attack the same leaf species. Thus, parasitoids in this lineage have evolutionarily labile host ranges, with little phylogenetic constraint on host shifts. This suggests that novel hosts entering the parasitoids' geographical range might be quickly adopted by one or more of the parasitoid species (Godfray et al. 1995). Similarly, if introduced into novel geographical areas, the parasitoids might readily use hosts that do not occur in their current home ranges (Henneman and Memmott 2001). However, leaf-mining insects are attacked by a particularly large number of parasitoid species of low specificity, compared with hosts in other feeding guilds (Hawkins and Lawton 1987), and we suspect that other host-parasitoid systems will show stronger signal through the parasitoid phylogeny.

We also tested whether two covariates, the geographical range of the host and the feeding mode of the parasitoid, affected association strength. The geographic range of the host was not a significant determinant of the number of parasitoid species that attack it. We might have expected an effect of the host geographic range if strict host specificity among parasitoids was common because such associations might have evolved more frequently if they involved widespread hosts (Hawkins and Lawton 1987). However, low specificity by parasitoids is the rule, making the lack of effect of host geographic range unsurprising. More surprising is the absence of a significant effect of endo- versus ectoparasitism because ectoparasitoids in general are known to have a broader host range than endoparasitoids (Askew and Shaw 1986; Sheehan and Hawkins 1991; Godfray 1994). Our analyses failed to reject the null hypothesis that ectoparasitoids attack more (or fewer) species of leaf miners.

Beyond the analyses of our specific example, the methods developed here can test hypotheses about general patterns of consumer-resource association while accounting for phylogenetic relationships. To give a concrete example, there has been much interest in the factors that determine the number of parasitoid species attacking different host species. Hawkins (1994) summarized a long series of studies and concluded that the manner in which the host species feeds on the host plant is the critical determinant, with leaf miners being attacked by the most species, followed by gall formers, species that live in webs and leaf rolls, external feeders, shoot borers, and finally root feeders. However, both feeding niche and the type of host attacked by parasitoids are strongly phylogenetically conserved. Thus, 80% of European leaf miner species occur in four monophyletic clades, and similarly, leaf miner parasitoids tend to come from a relatively small number of taxonomic groups (Askew 1980). There is almost certainly a very strong signal from both host and parasitoid phylogenies across the wide ranges of the hosts and parasitoids examined by Hawkins (1994), in contrast to our data set, which was limited to congeneric leaf miners. Unfortunately, the phylogenetic information is not available to test for phylogenetic signals in the broader collection of hosts and parasitoids, although this is likely to change in the next decade.

Our methods may also assist in applications surrounding intentionally or unintentionally invasive species. When a potential pest or harmful organism is recorded in a new area, it is of great importance to assess the risk of it going unchecked by natural enemies (Mooney and Drake 1986). Being able to predict from phylogenies which native consumers are likely to attack the new resource will help inform this risk assessment. Clearly, however, these predictions will need to be augmented by specific examination of the invading species' natural history and ecology (Williamson 1996). Similarly, biological control agents are frequently released to suppress nonnative pests, and there is growing concern about the possible side effects these releases may have on native organisms that are also attacked by the agent (Howarth 1991; Henneman and Memmott 2001). Careful risk assessments are now mandatory, and the techniques developed here may contribute toward a more objective assessment of the risks.

Acknowledgments

We thank T. Garland Jr., M. R. Helmus, and two anonymous reviewers for suggestions on the manuscript. This work was started as discussions at the National Science Foundation's National Center for Ecological Analysis and Synthesis, and it was supported in part by the National Science Foundation.

APPENDIX A

Estimating Ornstein-Uhlenbeck Process Parameters

Under the assumption of an Ornstein-Uhlenbeck (OU) process, the covariance matrix $\mathbf{V}(d_h)$ for hosts has elements $v_{ij}(d_h)$ given by

$$v_{ij}(d_h) = d_h^{v_{ii}+v_{jj}-2v_{ij}} \frac{1 - d_h^{2v_{ij}}}{1 - d_h^2}, \quad (\text{A1})$$

where values of v_{ij} on the right-hand side of equation (A1) are the covariances derived under the assumption of Brownian motion evolution (Blomberg et al. 2003, app. 2). A similar expression holds for the covariance matrix $\mathbf{U}(d_p)$ for parasitoids under the OU evolutionary process.

Parameters d_h and d_p can be estimated statistically by using estimated generalized least squares (EGLS). This method does not assume that trait values X_h and X_p are normally distributed or that the strengths of association (eq. [1]) are normally distributed. For the parasitoid attack rate A_{ik} , the EGLS estimates of d_h and d_p are obtained from equation (1) by constructing $\mathbf{W}(d_h, d_p)$ from $\mathbf{V}(d_h)$ and $\mathbf{U}(d_p)$. Letting $\mathbf{W}(d_h, d_p) = \sigma^2 \mathbf{C}(d_h, d_p)$, the EGLS estimates of d_h and d_p are those values that minimize the estimate of σ^2 (the mean squared error) under the constraint that $\det \mathbf{C}(d_h, d_p) = 1$. Specifically, for given values of d_h and d_p (Garland and Ives 2000),

$$\begin{aligned} \hat{b}_0 &= [\mathbf{I}_{1 \times n} \mathbf{C}(d_h, d_p)^{-1} \mathbf{I}_{1 \times n}]^{-1} [\mathbf{I}_{1 \times n} \mathbf{C}(d_h, d_p)^{-1} \mathbf{A}], \\ \hat{\sigma}^2(d_h, d_p) &= \frac{1}{N-1} (\mathbf{A} - \hat{b}_0)' \mathbf{C}(d_h, d_p)^{-1} (\mathbf{A} - \hat{b}_0). \end{aligned} \quad (\text{A2})$$

The value of $\hat{\sigma}^2(d_h, d_p)$ can be minimized numerically with respect to d_h and d_p . If \mathbf{A} is normally distributed, this procedure gives the maximum likelihood estimates of d_h and d_p (Judge et al. 1985). If \mathbf{A} is not normally distributed, the EGLS estimates of d_h and d_p will nonetheless be asymptotically unbiased, of minimum variance, and consistent (Judge et al. 1985).

To obtain statistical confidence bounds for d_h and d_p , we used a bootstrap procedure. Under the assumption that the model given by equation (1) is correct, standardized residuals can be computed as $\mathbf{D}(\mathbf{A} - \hat{b}_0)$, where \mathbf{D} is the matrix such that $\mathbf{D}\mathbf{C}\mathbf{D}' = \mathbf{I}$. Because $\mathbf{E}\{\mathbf{D}(\mathbf{A} - \hat{b}_0)[\mathbf{D}(\mathbf{A} - \hat{b}_0)]'\} = \sigma^2 \mathbf{D}\mathbf{C}\mathbf{D}' = \sigma^2 \mathbf{I}$, the standardized residuals all have the same variance, and the covariance between each pair of values is 0. Therefore, under the hypothesis that the fitted model is correct, the standardized residuals can be resampled (with replacement) to give bootstrap replication data sets with the same statistical

properties as the original data set. After fitting values of d_h and d_p using EGLS estimation, we produced bootstrap replication data sets from the standardized residuals, $\mathbf{D}(\mathbf{A} - \hat{\mathbf{b}}_0)$ and reassembled a data set by back-transforming over \mathbf{D} . We then fit each replication data set by using the same EGLS procedure that we used to fit the original data. Performing this procedure for 2,000 replication data sets gave 2,000 estimates of the paired values of d_h and d_p , and the joint frequency distribution of these values gives the bootstrap approximation of the probability distributions of the estimators of d_h and d_p from which we computed approximate confidence intervals (Efron and Tibshirani 1993).

A technical issue arises in our bootstrapping procedure. Although the procedure guarantees that the replicate residuals $\mathbf{D}(\mathbf{A} - \hat{\mathbf{b}}_0)$ have the same mean and variance as the standardized residuals from the original data, the back-transformed replicate data sets have qualitative differences from the original data. In particular, while the original data set contains many zeros (host-parasitoid pairs that do not interact), the replicate data sets do not. Similarly, when the binary measure of host-parasitoid association is used, the back-transformed data do not contain only zeros and ones. Thus, the bootstrapped data sets have different properties from the original data. Nonetheless, because the EGLS procedure is based on minimizing the variance among weighted residuals in the transformed space, the bootstrap procedure is likely to be insensitive to the distribution of data containing many zeros. Unfortunately, validating the bootstrap procedure is difficult because there is no way to generate simulation data sets having the same statistical properties of the original data. Therefore, we are forced to be cautious about the confidence intervals produced by our bootstrap procedure when many zeros occur in host-parasitoid associations or when using the binary measure of host-parasitoid association.

APPENDIX B

Predicted Associations for New Hosts or Parasitoids

Consider a host species z whose phylogenetic position is known. Suppose we seek to estimate the strength of the association A_{zk} between it and an arbitrary parasitoid k already present in the community. Let $\sigma^2\mathbf{C}$ denote the $N \times N$ covariance matrix fitted to the existing host and parasitoid species in the community. Let $\sigma^2\mathbf{C}_{zz}$ be the $m \times m$ matrix containing the covariances for the predicted strengths of association between the new host and each of the m parasitoid species in the community. The elements of $\sigma^2\mathbf{C}_{zz}$ are $v_{zz}u_{kl}$ ($k, l = 1, \dots, m$). Let $\sigma^2\mathbf{C}_z$ be the $m \times N$ matrix containing the covariances between the new host-parasitoid pairs and the remaining host-parasitoid

pairs for the new host; the elements of $\sigma^2\mathbf{C}_z$ are $v_{zj}u_{kl}$ ($j = 1, \dots, n; k, l = 1, \dots, m$). Denote by $\hat{\mathbf{A}}_z$ the $m \times 1$ vector of expected values of the parasitoid attack rates by the m parasitoids on species z . Under the assumption that the distribution of \mathbf{A} is multivariate normal (e.g., Fuller 1996),

$$\hat{\mathbf{A}}_z = \hat{\mathbf{b}}_0 - \mathbf{C}_z\mathbf{C}^{-1}(\mathbf{A} - \hat{\mathbf{b}}_0). \quad (\text{B1})$$

Furthermore, the estimate of the covariance matrix of the estimator $\hat{\mathbf{A}}_z$ is (e.g., Fuller 1996)

$$\hat{\Sigma}_z = \hat{\sigma}^2(\mathbf{C}_{zz} - \mathbf{C}_z\mathbf{C}^{-1}\mathbf{C}_z'), \quad (\text{B2})$$

from which standard errors of the estimates can be calculated. Although strict normality is unlikely to hold for real data sets, these results derived under the assumption of normality should provide reasonable approximations to the expected values of $\hat{\mathbf{A}}_z$ and their associated standard errors. This is demonstrated in figure 2.

Similar formulas can be derived in a straightforward way for associations involving a new species of parasitoid. Furthermore, partial information about a new species can be incorporated into the predicted parasitoid attack rates by including host-parasitoid pairs with known attack rates in \mathbf{A} (and matrix \mathbf{C}) and calculating $\hat{\mathbf{A}}_z$ for the unknown attack rates.

APPENDIX C

Alternative Definition of a_{ik}

For the analyses presented in the main text, we calculated the parasitoid attack rate a_{ik} by dividing A_{ik} by the average value of A_{ik} from all host species i from which parasitoid k was reared; this excluded zeros for host species from which parasitoids were not reared. Here, we give results for the alternative procedure for calculating a_{ik} by dividing A_{ik} by the average value of A_{ik} from all host species.

The alternative values of a_{ik} are positively correlated with the values used for the analyses in the text, although they are not extremely highly correlated (Pearson correlation, $r = 0.70$). Table C1 gives the results corresponding to table 1 for the alternative a_{ik} . The alternative a_{ik} gives a weaker phylogenetic signal from the host phylogeny (either the true phylogeny, \mathbf{V}_h , or the phylogeny constructed by grafting host species onto the tree phylogeny, \mathbf{V}_l) and a slightly stronger signal from the parasitoid phylogeny. The increase in importance of the parasitoid phylogenetic signal suggests that there is a signal in the number of host species attacked by parasitoids; the two procedures for calculating a_{ik} are sensitive to parasitoid host range because the al-

Table C1: Phylogenetic signal in parasitoid-host associations using the alternative procedure for calculating a_{ik} ; compare to table 1

Dependent variable, host phylogeny	d_h	d_p	MSE _d	MSE _{star}	MSE _b
$\sqrt{a_{ij}}$					
V_h	.086 (0, .24)	.21 (.042, .41)	.65	.70	.94
V_t	.35 (.19, .52)	.27 (.075, .44)	.60	.67	.77

ternative definition gives higher values of a_{ik} for parasitoid species that attack fewer species of hosts. However, this suggestion is refuted by the fact that there is no statistically significant parasitoid phylogenetic signal in the number of host species attacked per parasitoid (analysis not presented). Therefore, we are uncertain why the two procedures for calculating a_{ik} give different results, although the differences are minor.

APPENDIX D

Matlab Code for Table 1

The Matlab code that we used for the core analyses in the article (illustrated in table 1) may be found with this link: Matlab code. The code has not been peer reviewed, and neither the journal nor the authors are able to provide support. Programs for the other analyses in the article are available from A.R.I. on request. The MATLAB files provided include code (IG_AmNat.m [base code] and IG_funct.m [called by IG_AmNat.m]) and data files (hostV.txt, paraV.txt, plantV.txt, and hostparadata.txt).

Literature Cited

- Askew, R. R. 1980. The diversity of insect communities in leaf-mines and plant galls. *Journal of Animal Ecology* 49:817–829.
- Askew, R. R., and M. R. Shaw. 1986. Parasitoid communities: their size, structure and development. Pages 225–264 *in* J. K. Waage and D. Greathead, eds. *Insect parasitoids*. Academic Press, London.
- Blomberg, S. P., and T. Garland. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology* 15:899–910.
- Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.
- Charleston, M. A. 1998. Jungles: a new solution to the host/parasite phylogeny reconciliation problem. *Mathematical Biosciences* 149: 191–223.
- Combes, C. 2001. *Parasitism*. Chicago University Press, Chicago.
- Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall, New York.
- Emmet, A. M., I. A. Watkinson, and D. R. Wilson. 1985. Gracillariidae. Pages 244–363 *in* J. Heath and A. M. Emmet, eds. *The moths and butterflies of Great Britain and Ireland*. Vol. 2. Harley, Colchester.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1–15.
- . 2004. *Inferring phylogenies*. Sinauer, Sunderland, MA.
- Fuller, W. A. 1996. *Introduction to statistical time series*. Wiley, New York.
- Garland, T., and A. R. Ives. 2000. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *American Naturalist* 155:346–364.
- Garland, T., A. W. Dickerman, C. M. Janis, and J. A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. *Systematic Biology* 42:265–292.
- Garland, T. J., and R. Díaz-Uriarte. 1999. Polytomies and phylogenetically independent contrasts: an examination of the bounded degrees of freedom approach. *Systematic Biology* 48:547–558.
- Gauld, I. D., and B. Bolton. 1988. *The Hymenoptera*. Oxford University Press, Oxford.
- Gautier, N., J. Lasalle, D. L. J. Quicke, and H. C. J. Godfray. 2000. Phylogeny of Eulophidae (Hymenoptera: Chalcidoidea) with a reclassification of Eulophinae and the recognition that Elasmidae are derived eulophids. *Systematic Entomology* 25:521–539.
- Godfray, H. C. J. 1994. *Parasitoids: behavioral and evolutionary ecology*. Princeton University Press, Princeton, NJ.
- Godfray, H. C. J., D. J. L. Agassiz, D. R. Nash, and J. H. Lawton. 1995. The recruitment of parasitoid species to two invading herbivores. *Journal of Animal Ecology* 64:393–402.
- Grafen, A. 1989. The phylogenetic regression. *Transactions of the Royal Society of London B* 326:119–157.
- Hansen, T. E., and S. H. Orzack. 2005. Assessing current adaptive and phylogenetic inertia explanations of trait evolution: the need for controlled comparisons. *Evolution* 59:2063–2072.
- Harvey, P. H., and M. D. Pagel. 1991. *The comparative method in evolutionary biology*. Oxford University Press, Oxford.
- Hawkins, B. A. 1994. *Pattern and process in host-parasitoid interactions*. Cambridge University Press, Cambridge.
- Hawkins, B. A., and J. H. Lawton. 1987. Species richness for parasitoids of British phytophagous insects. *Nature* 326:417–423.
- Henneman, M. L., and J. Memmott. 2001. Infiltration of a Hawaiian community by introduced biological control agents. *Science* 293: 1314–1316.
- Housworth, E. A., and E. P. Martins. 2001. Random sampling of constrained phylogenies: conducting phylogenetic analyses when the phylogeny is partially known. *Systematic Biology* 50:628–639.
- Howarth, F. G. 1991. Environmental impacts of classical biological control. *Annual Review of Entomology* 36:485–509.
- Huelsenbeck, J. P., and B. Rannala. 2003. Detecting correlation between characters in a comparative analysis with uncertain phylogeny. *Evolution* 57:1237–1247.
- Ives, A. R., and J. Zhu. 2006. Statistics for correlated data: phylogenies, space, and time. *Ecological Applications* 16:20–33.
- Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T.-C. Lee. 1985. *The theory and practice of econometrics*. Wiley, New York.
- Lewis, O. T., J. Memmott, J. Lasalle, C. H. C. Lyal, C. Whitefoord, and H. C. J. Godfray. 2002. Structure of a diverse tropical forest

- insect-parasitoid community. *Journal of Animal Ecology* 71:855–873.
- Lopez-Vaamonde, C., H. C. J. Godfray, and J. M. Cook. 2003. Evolutionary dynamics of host-plant use in a genus of leaf-mining moths. *Evolution* 57:1804–1821.
- Martins, E. P. 2000. Adaptation and the comparative method. *Trends in Ecology & Evolution* 15:296–299.
- Martins, E. P., and T. F. Hansen. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *American Naturalist* 149:646–667 (erratum 153:448).
- MathWorks. 2005. Matlab, version 7.0. MathWorks, Natick, MA.
- Mooney, H. A., and J. A. Drake. 1986. Ecology of biological invasions of North America and Hawaii. Springer, New York.
- Müller, C. B., I. C. T. Adriaanse, R. Belshaw, and H. C. J. Godfray. 1999. The structure of an aphid-parasitoid community. *Journal of Animal Ecology* 68:346–370.
- Orzack, S. H., and E. Sober. 2001. Adaptation, phylogenetic inertia, and the method of controlled comparisons. Pages 45–63 in S. H. Orzack and E. Sober, eds. *Adaptationism and optimality*. Cambridge University Press, New York.
- Page, R. D. M. 1994. Parallel phylogenies: reconstructing the history of host-parasite assemblages. *Cladistics* 10:155–173.
- . 2000. TreeMap for Macintosh, version 2.0 beta. <http://www.cs.umd.edu/hcil/treemap/>.
- Page, R. D. M., P. L. M. Lee, S. A. Becher, R. Griffiths, and D. H. Clayton. 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. *Molecular Phylogenetics and Evolution* 9:276–293.
- Price, P. W. 1980. *Parasites*. Princeton University Press, Princeton, NJ.
- Purvis, A., and T. Garland. 1993. Polytomies in comparative analyses of continuous characters. *Systematic Biology* 42:569–575.
- Quicke, D. L. J. 1997. *Parasitic wasps*. Chapman & Hall, London.
- Rohlf, F. J. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* 55:2143–2160.
- Rothschild, M., and T. Clay. 1952. *Fleas, flukes and cuckoos*. Collins, London.
- Rott, A. S., and H. C. J. Godfray. 2000. The structure of a leafminer-parasitoid community. *Journal of Animal Ecology* 69:274–289.
- Sheehan, W., and B. H. Hawkins. 1991. Attack strategy as an indicator of host range in metopiine and pimpline Ichneumonidae (Hymenoptera). *Ecological Entomology* 16:129–131.
- Soltis, D. E., P. S. Soltis, M. W. Chase, M. E. Mort, D. C. Albach, M. Zanis, V. Savolainen, et al. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133:381–461.
- Strong, D. R., J. H. Lawton, and T. R. E. Southwood. 1984. *Insects on plants: community patterns and mechanisms*. Blackwood, Oxford.
- van Veen, F. J., R. J. Morris, and H. C. J. Godfray. 2006. Apparent competition, quantitative food webs and the structure of phytophagous insect communities. *Annual Review of Entomology* 51:187–208.
- Vinson, S. B. 1997. The behavior of parasitoids. Pages 417–469 in G. A. Kerkut and L. I. Gilbert, eds. *Comprehensive insect physiology, biochemistry and pharmacology*. Pergamon, New York.
- Williamson, M. 1996. *Biological invasions*. Chapman & Hall, London.

Associate Editor: Stephen B. Heard
Editor: Jonathan B. Losos