LETTER

The importance of plant relatedness for host utilization among phytophagous insects

Frode Ødegaard,1* Ola H. Diserud1 and Kjartan Østbye2
1Norwegian Institute for Nature Research, Tungasletta 2, NO-7485 Trondheim, Norway
2Department of Biology, University of Oslo, PO Box 1066, Blindern, NO-0316 Oslo, Norway
*Correspondence: E-mail: frode.odegaard@nina.no

Abstract
Current methods for measuring similarity among phytophagous insect communities fail to consider the phylogenetic relationship between host plants. We analysed this relation based on 3580 host observations of 1174 beetle species associated with 100 species of angiosperms in two different forest types in Panama. We quantified the significance of genetic distance as well as taxonomic rank among angiosperms in relation to species overlap in beetle assemblages. A logarithmic model describing the decrease in beetle species similarity between host-plant species of increasing phylogenetic distance explains 35% of the variation. Applied to taxonomic rank categories the results imply that except for the ancient branching of monocots from dicots, only adaptive radiations of plants on the family and genus level are important for host utilization among phytophagous beetles. These findings enable improvements in estimating host specificity and species richness through correction for phylogenetic relatedness between hosts and consideration of the host-specific fauna associated with monocots.

Keywords
Canopy crane, Coleoptera, evolution of host range, herbivore communities, host specificity, insect–plant interactions, Panama, plant taxonomy, species richness, tropical forests.

INTRODUCTION
Host relationships between insects and plants are one of the major factors driving evolutionary diversification within terrestrial ecosystems. Host specificity as the quantitative unit has received additional interest within the last few decades as a promising predictor of local as well as global patterns of biodiversity (Erwin 1982; Ødegaard 2000; Novotny et al. 2002a). In particular, significant progress has been made through the analyses of several extensive data sets of insect communities from tropical forests, in which the degree of effective specialization (May 1990), a widely used index for host specificity in insect communities, has been shown to be remarkably low, at between 5 and 10% (Basset et al. 1996; Ødegaard et al. 2000; Novotny et al. 2002a). Contrary to popular misconception (e.g. Erwin 1982), tropical forests appear to be dominated by generalist herbivores. Despite this, it is well recognized that most phytophagous insects tend to feed on more or less closely related host plants rather than a random selection of plants from a local flora (Eastop 1979; Strong et al. 1984). This is almost certainly because of evolutionary conservatism in host choice among insects, driven by the constraints of plant chemistry, taxonomic isolation and resource availability of plants (Southwood 1961; Lawton & Schröder 1977; Bernays & Chapman 1994). Consequently, the taxonomic relationship among host plants is likely to be a crucial parameter in studies of host specificity of insect assemblages, and especially in tropical forests where species rich genera and families dominate (Mawdsley & Stork 1997; Novotny et al. 2002a).

The connection between plant relatedness and host specificity can be expressed as the degree of similarity or dissimilarity between insect assemblages in relation to the phylogenetic distance between plants. Dissimilarity is closely linked to effective specialization (May 1990; Mawdsley & Stork 1997), whose value can be most readily conceptualized in the idealized case of host specificity in a sample with just two host species. In addition to general sampling effects, the magnitude of observed effective specialization will be affected by the number of host plant species and their taxonomic relatedness (Ødegaard et al. 2000). However, although the relationship between the number of plant species in the sample and effective specialization is well...
documented (Odegaard et al. 2000; Novotny et al. 2002a), the influence of the phylogenetic relationship between plants on the effective specialization of herbivores has never been properly quantified (Basset 1992; Novotny et al. 2002a).

It is of particular interest to quantify the relative importance of different taxonomic categories of plants for host utilization among phytophagous insects, because this is one of the key parameters in the estimation of global species richness (Erwin 1982; Odegaard 2000). For example, Novotny et al. (2002a) applied the number of plant genera as the taxonomic category of significance for host utilization among phytophagous insects. Mawdsley & Stork (1997) suggested that plant genera or even plant families may be the appropriate average taxonomic units for consideration of host specificity, while Odegaard (2000) employed the species level of plants combined with a between-community correction factor to achieve the same purpose. However, the problems with these approaches relate to missing or poor identification of phylogenetic distances between plants, and the shortage of large enough host-plant data for insects across a wide range of taxonomic relationships of plants.

In order to elucidate the relationship between the similarity of phytophagous insect communities between host plants and the degree of host-plant relatedness, we combined molecular data on phylogenetic distance between the host plants with one of the largest available data set on insect–host relationships, obtained from in situ canopy samples of phytophagous beetles associated with 100 plant species representing a wide range of taxonomic relationships of plants and two different types of tropical forests (Appendices S1 and S2).

METHODS

Sampling

The study was carried out at two Panamanian lowland forest sites situated close to the Atlantic Ocean and the Pacific Ocean entrances of the Panama Canal. Annual rainfall drops from 3100 to 1700 mm crossing the isthmus from north (Atlantic) to south and the forests at the two sites are classified as wet evergreen forest (San Lorenzo Protected Area), and deciduous dry forest (Parque Natural Metropolitano), respectively. A total of 1174 adult beetle species of Buprestisidae, Chrysomeloidea and Curculionoidea associated with a total of 50 and 52 plant species, respectively (Appendix S1), were sampled from canopy cranes in situ at 10–40 m height in the two forests. Potential pseudoreplication between sites was insignificant as only two plant species, Phryganocydia corymbosa (Vent.) Bur. and Cydista aequinoctalis (L.) Miess (Bignoniaceae), were shared between the two sites. A similar leaf and branch area of each plant was sampled by using a 1 m² beating sheet from the crane gondola. In order to estimate the variation within species, six individual trees of Brosimum utile (Kunth) Oken ex J. Presl (Moraceae) in the wet forest (Odegaard 2004), and two individual plants of 21 plant species in the dry forest were sampled. Sampling effort included c. 25 tree inspections per plant species, each of 30 min in duration, approximately every third week for 14 months in each forest (Odegaard 2005).

The total material included 65 831 beetles belonging to 2462 species from which 3582 host observations for 1550 beetle species were revealed by feeding records and host occurrences in situ according to Flowers & Janzen (1997) and Odegaard (2005). Host occurrences were assigned to the following abundance categories according to the number of individual records from the assumed host plants (number of records in brackets): 10 or more records (642); 5–9 records (554); 2–4 record (1214); 1 record and additional evidence for host association based of collections or literature (1172). In total, 568 (15.8%) of the host observations represented feeding records. Flower-visiting species were not included in the analyses.

Phylogenetic relationship

Reconstruction of phylogenetic relationships among higher taxa of plants has advanced rapidly in the last decade as accumulating DNA sequences provide new data sets (APG II 2003). The current classification is well established up to the rank of orders and a few monophyletic higher groups. In the present study the phylogenetic distance between pairwise combinations of plant species was estimated as the relative distance in number of taxonomic classification units between the two plant species, i.e. conspecific species: distance 1; congeneric species: 2; confamilial species: 3; shared order: 4; shared higher group ‘superorder’ (e.g. rosids, asterids, dillenids, magnoliids): 5; shared higher group ‘subclass’ (monocots, and eudicots): 6; shared class (angiosperms): 7. The classification follows APG II (2003) and assumes that taxonomic classification reflects phylogeny. The present rank of magnoliids as a part of a sister group of monocots is not clarified (Qiu et al. 1999; APG II 2003). Accordingly, we also tested magnoliids with ‘subclass’ rank (distance 6).

Sequences of the chloroplast gene rbcL were downloaded from GenBank (see accession numbers in Appendix S2), and aligned in SeqApp 1.9a157 (D. Gilbert, University of Indiana). The data set for the wet forest comprised 52 sequences (47 unique) of 1317 bp, with 126 and 318 characters being parsimony uninformative and informative, respectively, and 873 characters constant. For the dry forest, the data set comprised 50 (36 unique) sequences of 1402 bp, where 111 characters were parsimony uninformative, 365 were informative and 926 characters were constant.
Maximum likelihood (ML) analyses (heuristic search, 10 stepwise random taxon addition replicates, tree-bisection-reconnection, and confidence assessed by 100 bootstrap replicates) and minimum evolution (ME) analyses [maximum likelihood distance (D)], 10 stepwise random taxon addition replicates, tree-bisection-reconnection, and confidence assessed by 100 bootstrap replicates) were performed in Paup *4.0b10 (Swofford 1998) with a GTR + 1 + Γ DNA substitution model (Rodríguez et al. 1990) for both data sets as found in ModelTest 3.06 (Posada & Crandall 1998). We used ME estimates for comparison with relative taxonomic distance, and the best ML tree for a graphic display with bootstrap from both analyses.

For several of the studied plant species, DNA sequences did not exist, thus we had to rely on closely related species, and in some cases related genera, to assess phylogenetic relationships of plants in the two forests (Appendix S1). Thus, genetic distances at these taxonomic levels may include additional uncertainty. For congeneric pairs, we calculated an average genetic distance for all congeners of the study plants of which DNA sequences were available. This value \(D = 0.0089, n = 9\) was applied as an estimated genetic distance for congeners.

Statistical methods

Species overlap between beetle assemblages was measured as the Bray-Curtis similarity index (BC) for abundance data and the Sørensen coefficient (S) for presence/absence data (Magurran 2003). All the modelling and statistical analysis was performed with the free statistical software R (http://www.r-project.org).

With \(n\) host plants, we will have \(\binom{n}{2}\) \(\binom{n}{i}\) pair-wise similarities, but not all of these are independent. In order to remove the effect of covariance from the analysis, we must first identify which pairs of pair-wise similarities that are independent and which that are not. In total, we will have \(\binom{n}{2}\) pairs of pair-wise similarities. Two similarities will be independent if they are obtained from four different host plants, whereas dependence requires that they share one host plant. Dependent pairs of similarities will have a positive covariance. From the \(\binom{n}{2}\) \(\binom{n}{2}\) independent pairs of similarities, we can estimate the variance \(\sigma^2\) by the equation

\[
E[(BC_i - BC_j)^2] = 2\sigma^2,
\]

where \(BC_i\) is the similarity for host plant pair no. \(i\), and similarities \(i\) and \(j\) are independent. The covariance between the \(\binom{n}{3}\) dependent pairs of similarities can then be estimated from

\[
E[BC_k - BC_l] = 2\sigma^2 - 2\text{Cov},
\]

where the host plant pairs \(k\) and \(l\) share one host plant. The effect of this covariance on the analysis will decrease with increasing number of host plants in the data set. With \(n = 3\), we will only have dependent pairs of similarities, but with \(n = 50\), less than 8% of the pairs of similarities will be dependent. In all of our subsequent statistical analysis, the effect of this covariance has been removed and only the variances estimated from the independent pairs of similarities are applied.

The modelling of the relation between similarity (BC) and genetic distance (D) was carried out by fitting a nonlinear regression model. A log-model seemed reasonable as we expected a sharp (exponential) increase in similarity for host plants with very short genetic distances. However, because of individual variation and sampling effects (Ødegaard 2004), similarity for conspecific plants did not exceed 60%. The log-models explained about twice as much of the total variance (twice as large \(R^2\)) as simple linear models did.

The study of similarity vs. taxonomic rank categories was carried out as pairwise comparisons of mean beetle similarities between two consecutive taxonomic rank levels, i.e. six ordinary two-sample \(t\)-tests. This approach highlights the difference in similarity between neighbouring categories.

RESULTS

Similarity (BC) in beetle assemblages decreased with genetic distance (D) between the host plants (Fig. 1). Plant relationship explained 40% and 33% of the variation in similarity of beetle communities in San Lorenzo Protected Area (wet forest) (BC = -10.7 – 8.1 log(D); \(R^2 = 0.40\) \(P < 0.001\)) and the Parque Natural Metropolitano (dry forest) (BC = -4.3 – 6.9 log(D); \(R^2 = 0.33\) \(P < 0.001\)), respectively. The models were significantly different (\(t\)-test for comparison of slopes for the two models: \(P < 0.001\)). In order to achieve an operational response curve for use across other tropical forests, a combined model for both forests was calculated by pooling the similarities and genetic distances from the two forests together, and fitting a model to the pooled data set. We have not calculated similarities or genetic distances across forests, as the beetle-species overlap between the two forests was negligible (Ødegaard 2005). At the least, this intermediate model \([BC = -7.4–7.4 \log(D); R^2 = 0.55\) \(P < 0.001\]) can be applied to forest types within the range of the two presented here.

With regard to taxonomic categories, there was a steeper drop in beetle similarity when phylogenetic distance was small (between congeneric or confamilial plants) and large (between plants belonging to different subclasses) (Fig. 2). On intermediate taxonomic scales, i.e. from family to subclass, there were only small changes in similarity across taxonomic categories. On average, for both forests, the similarity decreased 50.0% from species to genus level;
69.4% from species to family level; 77.7% to the level of order; 80.8% to the level of superorder; 82.0% to the level of subclass, and 94.6% to class level. When we performed two sample t-tests to compare mean similarity between consecutive taxonomic categories, we found for the wet forest that all neighbouring categories had significantly different similarities ($P < 0.01$), except from family to order ($P = 0.25$). For the dry forest, the difference in mean similarity between consecutive taxonomic categories was significant ($P < 0.01$), except from genus to family ($P = 0.054$), from order to superorder ($P = 0.80$) and from superorder to subclass ($P = 0.73$). Different mean similarities for the taxonomic categories explained 47% and 40% of the variation in beetle similarity for the wet and dry forests, respectively. Piecewise linear curves between means of consecutive taxonomic categories illustrates how the beetle similarity decrease with higher taxonomic rank categories (Fig. 2).

When applying the widely used Sorensen coefficient on presence/absence data, similar models were found both for genetic and taxonomic distances although explained variance decreased, i.e. the models had lower $R^2$ values.

**DISCUSSION**

**Evolutionary implications**

These results have theoretical implications for the evolution of host range and host specificity of phytophagous insects. The evolutionary divergence of monocots from dicots in the Early Cretaceous 120 million years ago (Wolfe et al. 1989) was clearly important from the perspective of host – herbivore relationships as few phytophagous insects are shared between monocots and other angiosperms (Fig. 2). The result from the present study was based on only three species of palms, but there is no reason to doubt its general validity as many higher groups of phytophagous insects are mainly restricted to monocots, e.g. Criocerinae, Donacinae, Cephaloleiini (Hispinae) within Chrysomelidae (Jolivet &
Hawkeswood 1995), and Cholinae, Derelomini and basal groups within the Curculionidae (Marvaldi et al. 2002). The placement of monocots in the root of the angiosperm phylogenetic tree is well supported (Qiu et al. 1999), and their unique position as a relatively old isolated and species-rich group of angiosperms may explain their status as the only higher ranking angiosperm-taxon that matters to the beetles. The non-quantified phylogenetic distance between different taxonomic rank categories (Fig. 2), therefore, does not reflect the true genetic distance between the monocot palms and the other angiosperms (Appendix S1), which may explain the sigmoidal shape of the curves (Fig. 2). However, particular ecological characters in monocots, not reflected by genetic distance, may be selected by the beetles.

Magnoliids, which are classified as part of a sister group to monocots (APG II 2003) did not have beetle faunas significantly different from superordinal groups within eudicots (Mann–Whitney test, wet forest: \( z = 0.381, P = 0.703 \); dry forest: \( z = 1.707, P = 0.088 \)). Accordingly, beetle host-plant preferences interpreted as a taxonomic character support the classification of magnoliids as a superorder closer to eudicots. The diversification of angiosperms into superordinal, and ordinal ranks did not affect similarity in beetle assemblages, which corresponds with an observed lack of insect taxa associated with host plants across a broad array of subgroups within any of these higher taxonomic categories of angiosperms. Thus, the broad generalists among phytophagous beetles may have switched across plant groups of intermediate taxonomic rank several times during their evolutionary history. The observation that old groups of insects are associated with old groups of plants (Ehrlich & Raven 1964; Farrell 1998) should therefore be restricted to higher groups of plants like cycads, conifers, monocots and dicots as indicated by Marvaldi et al. (2002).

Our results confirm that many beetles are specialized within plant families and genera (e.g. Strong et al. 1984). This corresponds with the hypothesis that the phytophagous beetle lineages represent a series of adaptive radiations onto angiosperms, with subsequent conservatism in the evolution of beetle–plant associations that maintain specialization (Farrell 1998). A similar pattern of host use was found for Lepidoptera (Novotny et al. 2002b), suggesting a general validity of the phenomenon for herbivorous insects. The similarity drops at decreasing rates from species to genus, to family, and to ordinal rank, which indicates that the most recent taxonomic branching events are most important for host utilization among phytophagous beetles.

**Estimation of host specificity and species richness**

The floral composition of a defined area is the baseline for the degree of specialization among herbivores in the local community. The data suggests that the genus, family and subclass levels of angiosperms are particularly important for evolution of dissimilarity and host specificity in phytophagous insect communities. Therefore, in order to estimate host specificity in an insect community, the phylogenetic distances between the studied host plants should be similar to that of the hosts in a larger area, at least for the latter taxonomic categories. Alternatively, the ratio between congenerics, confamilials and species that share classes of angiosperms can be corrected according to the present model, assuming the studied forests are representative of tropical forests in general. Knowing the range of insect host-plant species, and the ratio between taxonomic categories among the hosts, future studies will be able to use a host-relationship coefficient that will correct bias in calculation of average similarity in insect assemblages caused by taxonomic relatedness between plants. Hence, future studies on host specificity may be less time- and resource-intensive as fewer host species will need to be surveyed.

Ecological estimates of local as well as global species richness of arthropods (Erwin 1982; Ødegaard 2000; Novotny et al. 2002a) may consequently be improved by correcting for phylogenetic relatedness between hosts according to the present model. For instance, the recognition that beetle similarity on average drops 50 and 70% from species to genus and family level of plants, respectively (Fig. 2), suggests a significant increase in estimated species richness for calculations based on these higher taxonomic categories of plants as the average units for host utilization among phytophagous insects. Moreover, ecological estimates of species richness are mainly based on assessments of host specificity in insect assemblages associated with trees among eudicot angiosperms, despite the fact that monocots make up 24% of all angiosperms (Heywood 1993). Certainly, most monocots are small plants with comparatively poor herbivore faunas, such as Graminaceae and Orchidaceae (Lawton & Schröder 1977), but consideration of palms, and larger shrubs and epiphytes among the monocots would probably increase global diversity estimates as their insect faunas are species rich and at the same time very different from those of eudicots (Memmott et al. 1994; Wilf et al. 2000; Howard et al. 2001). Combined with new insect data on the degree of beta diversity (Ødegaard 2005) and biased sampling effects (Ødegaard 2004) in tropical forests, global species richness should probably remain at the upper end of the 5–10 million species interval (Ødegaard 2000).

**ACKNOWLEDGEMENTS**

We thank the Canopy Crane team at STRI, and several taxonomists acknowledged elsewhere; C. L. Nesbø and B. H. Wilmann for assistance with phylogenetic and...
statistical analyses; and D. Halley, V. Novotný, O. T. Sandlund, A. Viken and two anonymous referees for valuable comments on the manuscript. This work was financially supported by the Norwegian Research Council, the Smithsonian Institution Fellowship Programme, and the Norwegian Institute for Nature Research.

SUPPLEMENTARY MATERIAL

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/ELE/ELE758/ELE758sm.htm

Appendix S1 Phylogenetic relationship between host plants in the wet forest and the dry forest as given by the best maximum likelihood tree.

Appendix S2 The plant species studied and their substitution number in the wet and dry forest.

REFERENCES


Editor, Mark Schwartz

Manuscript received 18 January 2005
First decision made 12 February 2005
Manuscript accepted 04 March 2005