

Fernley B. Symons · George W. Beccaloni

## Phylogenetic indices for measuring the diet breadths of phytophagous insects

Received: 16 November 1998 / Accepted: 10 February 1999

**Abstract** Prevailing methods of measuring diet breadth of phytophagous insects are not consistent between studies and generally rely on counts of a variety of higher plant taxa (e.g. genera, families, orders). Results derived from them can be inconsistent if different taxonomic levels are used between studies. In any case, such indices do not include information from the whole branching structure of the host plant phylogeny, and do not address the fact that higher taxa are not necessarily phylogenetically equivalent. Here we present novel phylogeny-based methods which address these shortcomings. Although a previously proposed index (the Phylogenetic Diversity index) may be employed, it cannot be used to measure diets of strictly monophagous insects (i.e. those which utilise a single host species). We therefore introduce a modification of this index (the Root Phylogenetic Diversity index) which may be applied to all diets. In addition, we propose a Clade Dispersion index as a branch-length-independent measure of the degree to which hosts are scattered across the host phylogeny. We describe how these indices could be employed in studies of insect diet breadth and discuss potential problems which may be encountered in their use.

**Key words** Host plants · Minimum spanning subtree · Phylogenetic Diversity index

### Introduction

A fundamental concern of ecologists is to understand the determinants of the relative size of an organism's

niche (e.g. MacArthur 1972; Mitter and Farrell 1991). Studies of diet have been important in this regard, as this niche dimension easily lends itself to quantitative analysis. The diets of phytophagous insects in particular have been extensively studied. There are a number of reasons for this: these species and their host plants together represent a large proportion of terrestrial biodiversity (Strong et al. 1984); in many cases, insects may be reared in captivity in large numbers, thereby allowing the assessment of preferences and determination of nutritional ecology; and there is much variation in the number and identity of hosts consumed, which provides abundant material for the statistical investigation of macroecological or evolutionary patterns. Larval development is often completed on a single individual of a host species, ensuring that there is a close correspondence between other dimensions of the insect's niche and its diet breadth (see also Dethier 1954; larval stages usually comprise most of the life cycle). Here, we give a brief overview of the ways in which the diets of phytophagous insects have been measured in the literature and propose novel methods which avoid some of the problems associated with these approaches. Although we restrict our discussion to phytophagous insects, the methods presented are of general utility and could easily be adapted for use with other groups of organisms.

### Prevailing methods of diet breadth measurement

Despite the large number of studies of insect diet breadth, across-study comparison is hampered by the fact that authors often use incompatible diet breadth indices (see discussions in Jaenike 1990; Schoonhoven et al. 1998). Some examples of these are presented in Table 1. Similarly, different definitions of the terms monophagous, oligophagous, stenophagous, euryphagous and polyphagous appear in the literature. In general, diet breadth has been measured as the number of host taxa consumed at a particular level of the host

F.B. Symons  
Department of Zoology, University of Oxford,  
Oxford, OX1 3PS, UK

G.W. Beccaloni (✉)  
Department of Entomology,  
The Natural History Museum,  
Cromwell Road,  
London, SW7 5BD, UK  
e-mail: g.beccaloni@nhm.ac.uk, Fax: +44-171-9388937

**Table 1** A selection of diet breadth measures which have been employed in studies of phytophagous insects and their hosts. *Rank* of division is as given in the study; we have inferred ranks where these were not given explicitly

Reference	Rank	Diet breadth categories	Name of division (if given)
Joern (1979)		Shannon Information index with host species <sup>a</sup>	
Cates (1981)	1	Number of host species ≥1 host in 1 genus	Monophagous
	2	≥2 host genera in 1 family or closely related families	Oligophagous
Hayes (1982)	3	Hosts in ≥2 families	Polyphagous
	1	Hosts in 1 genus	Monophagous
	2	Hosts in 1 order, > 1 genus	Oligophagous
Niemelä et al. (1982)	3	Hosts in > 1 order	Polyphagous
	1	Hosts in 1 family	Specialists
	2	Hosts in > 1 family	Generalists
Brown and Southwood (1983)	1	1–4 hosts in 1 genus	
	1.5	≥5 hosts in 1 genus	
	2	2–4 host genera in 1 family	
	2.5	≥5 host genera in 1 family	
	3	> 1 host family in 1 order	
	4	Hosts in ≥2 orders in 1 super order	
Gaston and Reavey (1989)	5	Hosts in ≥2 super orders	
	1	Hosts of 1 species	
	2	Hosts in 1 genus	
	3	Hosts in 1 family, > 1 genus	
Dyer and Floyd (1993)	4	Hosts in > 1 family	
	1	Hosts in < 2 families or hosts of 1 species	Specialists
Novotný (1994)	2	Hosts in > 5 families	Generalists
	1	Hosts in 1 order	Specialists
Fiedler (1995)	2	Hosts in 2–3 orders	Polyphagous
	1	Hosts in 1 genus	
	2	Hosts in 1 family	
	3	Hosts in 2 families	
	4	Hosts in 3 families	
	5	Hosts in 4–5 families	
Fiedler (1998)	6	Hosts in > 5 families	
	1	Hosts in 1 family	Monophagous
	2	Hosts in 2–3 families	Oligophagous
	3	Hosts in ≥ 4 families	Polyphagous
	1	Hosts in 1 family	Monophagous
2	Hosts in > 1 family	Polyphagous	

<sup>a</sup> Includes information about relative abundances of diet components

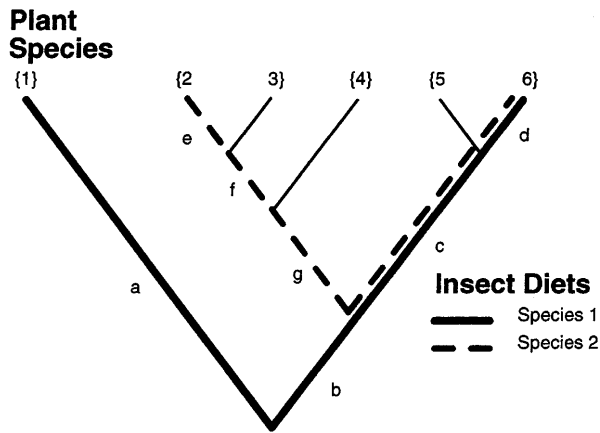
taxonomy, although more sophisticated indices which employ some estimate of host preference have also been used (Table 1). Adoption of a common measurement methodology such as that proposed below would help to clarify the situation.

### Problems associated with counts of higher host taxa

There are a number of problems associated with prevailing methods for quantifying diet breadth. Counting the numbers of host taxa utilised at one particular taxonomic level (e.g. counting the number of host families utilised) can lead to contradictory results when compared with results from counting host taxa utilised at other taxonomic levels (e.g. counting the number of host genera utilised). Thus, in a study of the hosts of the North American and Australian butterfly faunas (G.W. Beccaloni and F.B. Symons, unpublished data), different results were obtained depending on the level of the hosts' taxonomy at which diet breadth was determined. This

type of inconsistency is the primary reason why studies of diet breadths cannot always be compared. Additional problems associated with not taking the whole branching structure of the hosts' phylogeny into account are examined below.

Two host higher taxa (e.g. two plant families) are not phylogenetically equivalent entities unless they happen to be sister taxa. Thus, counts of the numbers of host higher taxa utilised generally fail to take into account the degree of relatedness of the counted taxa. For example, in Fig. 1 both insect species utilise plants from two families, but the diet of insect species 1 is arguably broader than that of insect 2, as the families it feeds on are less closely related to one another than are those utilised by insect 2. These insects would have equivalent diet breadths according to many prevailing diet indices. Unfortunately, other more sophisticated diet breadth measures which include information about host plant preferences also treat host taxa as if they are phylogenetically equivalent (e.g. the Shannon Information index as employed by Joern 1979).



**Fig. 1** A hypothetical phylogeny of six plant species showing the shortest pathways between the hosts of each of two insect species. These are used in calculating Phylogenetic Diversity index values (see text). Lower case letters identify branch segments included in these pathways. Plant families are designated by brackets (e.g. {2 3}). Note that branch lengths are shown as ultrametric for convenience only. Real branches might be of any length

## Phylogenetic indices to measure diet breadth

### The Phylogenetic Diversity index

The Phylogenetic Diversity (PD) index has been widely implemented in the conservation literature as a way of estimating biodiversity at threatened sites (see Faith 1992a, 1992b, 1994; Williams et al. 1994). It is defined as the sum of the branch lengths of the minimum inclusive clade (the minimum spanning subtree) for the organisms under consideration. It therefore can only be calculated for diets comprising two or more species. For example, in Fig. 1, the PD value for the diet of insect species 1 would be the sum of the branch lengths  $a + b + c + d$ , while the PD value for the diet of species 2 would be the sum of the branch lengths  $c + d + e + f + g$ . Where the aim is to quantify diet breadth with respect to plant phylogeny, then branch lengths should reflect overall genetic distances between plant species.

In contrast to some prevailing methods of measuring diet breadth, the information used to calculate the PD will not necessarily come from a single taxonomic level alone. Since the overall topology of the plant phylogeny defines the minimum inclusive clade for a particular diet, information from the whole of the plant phylogeny is included in the calculation of the index. Standardising the index (see below) incorporates even more information about the overall host phylogeny.

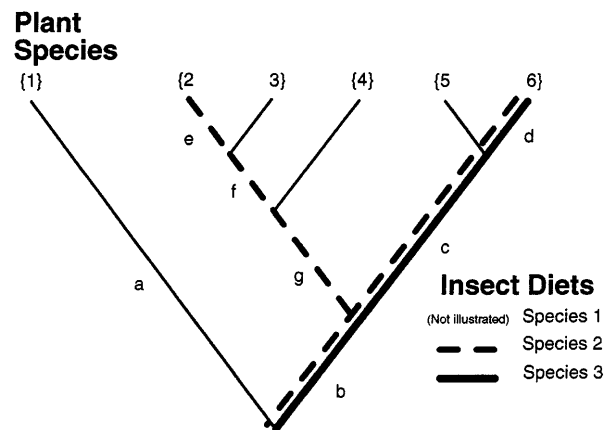
### The Root Phylogenetic Diversity index

As discussed above, PD values cannot be calculated for the diets of strictly monophagous insects (i.e. those which utilise a single species of host plant). However, a

modification of the PD index provides a solution to this problem. We term this the Root Phylogenetic Diversity (Root PD) index. The Root PD for diets which include two or more host species is defined as the PD plus the sum of any branches along the shortest path from the basal node of the minimum inclusive clade to the root-most node of the phylogeny (where a diet does not span the root of the tree: e.g. insect species 2 in Fig. 2). In the same way, the Root PD value for a monophage is the sum of the branch segments along the shortest path from the tip representing its host species to the root-most node of the host phylogeny. For example, the Root PD for the insect diets shown in Fig. 2 is the sum of branches  $b + c + d + e + f + g$  for insect species 2, and the sum of branches  $b + c + d$  for species 3 (a monophage). This index has the additional advantage that pair-wise phylogenetic similarities may be calculated for all diets, even those which encompass a single host species (these estimate the phylogenetic covariances of diets when character evolution approximates a Brownian motion model). Plainly, the assignment of a root node for Root PD calculation is an arbitrary procedure. It is not meaningful to directly compare Root PD values between studies where different roots have been designated.

### The Clade Dispersion index

We introduce the Clade Dispersion (CD) index as a kind of nearest-neighbour index for host plants. It is obtained by dividing the observed number of branches separating taxa in a diet by the minimum number of branches possible for that number of taxa given the branching structure of the phylogeny. In a fully bifurcating phylogeny, the minimum number of branches separating taxa can be estimated as  $(2n - 2)$ , where  $n$  is the number of host taxa. For example, in Fig. 1, both insect species 1 and 2 utilise two host species and therefore  $n = 2$ ,



**Fig. 2** A hypothetical plant phylogeny as in Fig. 1. Pathways used for the calculation of Root Phylogenetic Diversity index values are shown for the diets of two insect species. Lower case letters identify branch segments included in these pathways

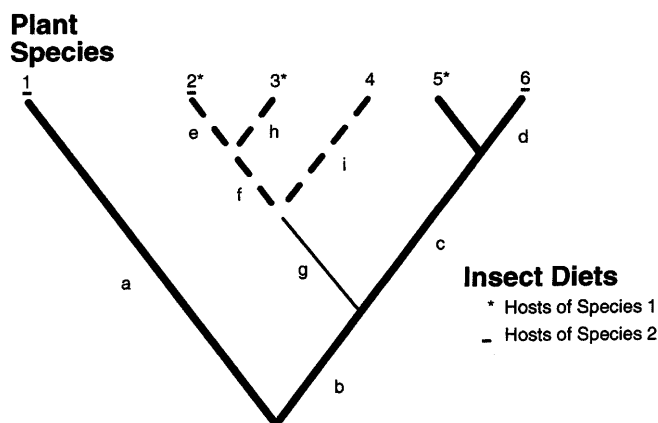
giving a minimum path estimate of 2 in both cases. Thus the CD index value for species 1 in Fig. 1 is  $4/2 = 2.0$  and for species 2 it is  $5/2 = 2.5$ . Use of the minimum-path estimate tends to overestimate the CD index in real phylogenies, however, because intervening branch segments may cause the smallest number of path segments to be greater than  $2n - 2$ . For example, in Fig. 1, the minimum number of branch segments that have to be traversed to encompass the diet of an insect which utilises plant species 2, 3 and 5 is six segments, two more than the naive ( $2n - 2$ ) estimate. A possible way of removing such confounding effects of branching structure might be to proceed as follows. Using each host species in a diet as a starting point in turn, one would estimate the minimum number of branch segments separating the number of taxa in the diet. The mean of these values could then be used as the denominator for the observed number of segments traversed. In Fig. 3, we present an example of this methodology for an insect with three hosts (insect species 1 which utilises host species 2, 3, and 5). Assuming that these three hosts could be as near to each other as possible on the plant phylogeny, the minimum possible number of branches which encompass these hosts would be four branches starting from either host 2 or 3 (dotted line), and five branches starting from host 5 (bold line). (Note that an alternative minimum pathway to that illustrated exists for this latter host species. This would incorporate branches g and i in place of branches a and b.) After calculating such minimum pathways, their mean can be used as the denominator in calculating CD index values. For the insects in Fig. 3, average pathway values would be  $(4 + 4 + 5)/3 = 4.33$  for insect species 1, and  $(5 + 5 + 4)/3 = 4.67$  for insect species 2. The CD index value for insect species 1 would therefore be  $6/4.33 = 1.39$  and for insect 2 it would be  $7/4.67 = 1.5$ . This pathway method makes no distinction between diets where most host species are nearest neighbours but there is a single or a few disjunct

hosts (e.g. the diet of insect 1 in Fig. 3), and those where all host species are disjunct (e.g. the diet of insect 2 in Fig. 3).

The CD index as defined above implicitly assumes that host phylogeny branch lengths are equal since all branch segments are treated alike. Since this index estimates the degree to which an insect is allied to a particular clade of hosts (as defined either according to host phylogeny or according to insect perceptions of those hosts; see below), it is most likely to be of use in coevolutionary studies or in investigations of adaptive radiations. Note that as in the case of the PD index, CD values cannot be calculated for strictly monophagous species. Where the comparison of distributions of CD (or PD) values is important, CD values for the diets of strict monophages could be set to some arbitrary constant.

### Modifications of the phylogenetic diversity indices

Both versions of the PD discussed above confound the effects of the relative relatedness of taxa and the number of species in the diet. Thus, the relative relatedness of host species may be estimated by dividing either PD index by the number of species utilised. We propose to call the PD index thus modified the Per-Taxon Phylogenetic Diversity (PTPD) index. Where diet breadths are being compared between insect species, the PD (and/or Root PD), the PTPD (and/or Per-Taxon Root PD), the CD index, and counts of numbers of species utilised in a diet can be used together to determine, first, whether a diet differs from other diets, and second, to assess what aspect of the diet causes these differences. Note that it is possible for numbers of species utilised to differ significantly between diets, whilst the values of the indices given here may not differ significantly. The explanation is that the present indices include information about the relatedness of those species which is not present in simple species counts. A diet which consists of several very closely related species is arguably less broad than one which consists of slightly fewer distantly related ones.



**Fig. 3** A hypothetical phylogeny with ultrametric branches of six plant species. The hosts which make up the diets of two insect species are indicated in the diagram. Pathways encompassing minimum numbers of branches for hosts in the diet of insect species 1 are indicated (these are not given for the hosts of insect 2)

### Problems and prospects

#### Comparisons of diet breadths between regions

The fact that host plant species are likely to be restricted to a region of study makes comparing diet breadth index values between regions difficult. Plainly, diets may only include those plants which occur in the same places as the study insects (e.g. Dethier 1954). This means that differences in the values of those diet breadth indices which utilise branch length information explicitly (such as the PD) may in fact be due to biases in the flora of a given region. This is because the particular values for an index obtained in one region may simply not be possible given the branching structure of the host plant phylogeny for other regions where those hosts do not occur.

Note that if one had access to the true phylogeny (with true branch lengths) for the hosts of each region, direct between-region comparisons could be made between the indices proposed above. Even here, however, the potential problem of differing phylogenies between regions not allowing the same distribution of possible diet breadths would mean that it would not be possible to identify the causes of any between-region differences in diet breadth. We discuss a partial solution to this problem below. We stress that it is even less meaningful to make between-region comparisons of diets which have been measured by prevailing methods. For example, comparison of counts of host plant families utilised tell the investigator little about relative diet breadths if the families counted are not phylogenetically equivalent and/or differ between regions.

A partial solution to the problem of across-region comparisons might be to standardise diet breadth index values. Each diet breadth index value for a particular region could be divided by the sum of all the branch lengths in the regional host plant phylogeny and these standardised values could then be compared across the regions of interest. Where the host phylogenies from two regions are very large, the range of possible standardised diet breadth index values is likely to be fairly similar for both areas, and so diet breadth indices will more accurately reflect any quantitative differences in host plant usage.

Phylogeny estimates for host plants are dependent on the numbers and identities of the taxa which are employed in their reconstruction (Lecointre et al. 1993). To ensure that the branching structure of the reconstructed host phylogeny is consistent between regions/studies, a composite phylogeny based on all species in the combined host lists of the regions/studies in question can be constructed. Note, however, that CD index calculation can be biased by indiscriminate use of composite phylogenies. This is because diets encompassing host species that are as closely related as is possible given a particular local flora could appear disparate (have large CD values) when extralimital plant species are included, if the latter species cause an increase in the observed number of branch segments required to encompass a diet. It is therefore probably better to exclude those taxa which are extralimital to the region of interest from the composite phylogeny before calculating CD values. This would retain the overall branching structure of the composite phylogeny but may, in some cases, reduce the number of branch segments separating taxa.

The preceding discussion emphasises that underlying phylogenetic biases in the composition of regional floras may influence diet breadth over and above factors directly affecting insect ecology. This fact has rarely been addressed in macroecological studies and is distinct from biases caused by differential abundances of host species (reviewed in Strong et al. 1984). Macroecological studies which involve host plant phylogenies generally assume that all insect species have equal access to all plant species, although partitioning by, for example, habitat

means that this assumption is probably unrealistic in most cases (note, however, that changes in climate might have acted to make currently unavailable hosts accessible to insect species).

#### Diet breadth with respect to what?

There is increasing evidence that insects use a single trait or a suite of a few traits to classify host plants (reviewed in Jaenike 1990). The distribution of these traits may not be in accordance with the overall phylogeny of the hosts (see for example Futuyma 1983; Mitter and Farrell 1991). Indeed, it is often not clear which taxonomic level best reflects changes in the particular trait(s) of interest even in those cases where the distribution of traits *is* well predicted by overall plant phylogeny. This is not surprising considering that host plant taxonomic designations are generally made with reference to many different characters, no one of which may map exactly onto the resultant taxonomy. Despite this, prevailing methods of measuring diet breadth assume that counts of higher taxa are valid surrogates for the number of different combinations of host discrimination traits which are encountered by insects. In contrast, we argue that a diet breadth index which employs information from all levels of the host's phylogeny is preferable where there is no direct knowledge about the distributions of the traits used in host selection, since overall phylogeny is the best overall predictor of the distribution of traits (see for example Farris 1979; Mitter and Farrell 1991).

The fact that certain host traits may be more important to the insect than others highlights a significant point: is the aim of a given study to quantify diet breadth according to the overall similarity of the host plants, or is it to quantify diet breadth with respect to that subset of plant traits which insects use to distinguish host plants? A study of the latter type can be considered as an attempt to quantify diet breadth from the insect's perspective. We argue below that studies of diet breadth with respect to overall plant phylogeny should be seen as the logical precursors to studies which attempt to assess diet breadth from the insect's perspective.

Given that plant phylogeny represents the best overall estimate of similarity between plant taxa, diet breadths measured according to some other criterion/criteria which is/are thought to be important in discrimination by insects (e.g. the similarity of secondary chemistry of host plants) may only be assessed in the light of diets measured on the overall plant phylogeny. Thus, the fact that *Pieris brassicae* L. utilises hosts from the families Brassicaceae, Tropaeoleaceae and Resedaceae (Schoonhoven et al. 1998) is of interest because these taxa are not closely related and are dissimilar in many respects. However, all synthesise glucosinolates which adult *P. brassicae* use to discriminate hosts from non-hosts.

When quantifying diet breadth from an insect's perspective, a phylogeny estimate should be constructed

using just those host characters which are thought to be of (primary) importance to the insects (e.g. see Futuyma and McCafferty 1990; Becerra 1997; Janz and Nylin 1998). Note that in the studies by Futuyma and McCafferty (1990) and Becerra (1997), phenograms rather than phylogeny estimates of host relationships were constructed using secondary chemical information. We advocate the use of phylogeny estimates based on a reduced set of characters rather than phenograms, on the basis that the latter may produce trees which differ extensively according to the algorithm used (see Farris 1979). Hereafter we refer to this type of 'reduced' phylogeny estimate as a dendrogram to distinguish it from estimates of host phylogeny which use all available characters.

There is the potential to use the results from preference testing as 'characters' for the construction of such dendrograms. The rationale for so doing is that the insects are responding to traits or suites of traits of the host plants and so preferences simply reflect the underlying distribution of (unknown) traits amongst the plants. This methodology makes no coevolutionary assumptions about the relationship between the insects and their hosts. However, it is likely that such characters will be non-independent, since closely related insects are likely to use similar host traits as cues (Harvey and Pagel 1991). Nevertheless, the dendrogram method is likely to be most useful when investigating the diet breadths of single insect species or those of closely related species. To investigate larger assemblages, one has to assume that the characters chosen for dendrogram construction are those utilised by all the insect species studied.

Comparison of diet breadths measured on the overall plant phylogeny with those measured on a dendrogram or dendrograms may be informative (see below). In those cases where insects appear to be using host cues which are congruent with host phylogeny or where insect and host phylogenies are congruent (e.g. Farrell and Mitter 1990), application of the indices proposed earlier is straightforward.

It can be important to incorporate information about varying degrees of host usage or differing preferences for different hosts. This may be achieved by using a branch length weighting scheme for the host phylogeny or dendrogram.

### Investigating host selection

A diet measured on the plant phylogeny can be seen as a sort of 'null value', as it estimates the diet breadth with reference to the overall distribution of host characters. Diet breadths obtained from dendrograms constructed as discussed above may be compared with these null diet breadth values in order to investigate which characters are important in host selection. Both dendrogram diet breadths and null diet breadths must be standardised (see above) prior to such comparisons.

We envisage that standardised diet breadths measured on a dendrogram will frequently be smaller than the null values where insects mainly use one or a few host traits in making host plant choices. However, this may not be the case for all insects [e.g. some Orthoptera which select hosts according to the principle of nutrient complementarity (Bernays 1998): here diets may appear broader when measured on a dendrogram]. Crude comparisons between hypotheses regarding which suites of host traits are important in discrimination could be made by comparing diets measured on dendrograms constructed from different sets of traits. However, this method suffers from the problem discussed earlier, namely that the exact branching structure of the dendrograms determines the values which a diet index can attain and so diet breadths may not be exactly comparable (especially where the dendrograms are very different).

### Null-model approaches

A null-model approach could be used to investigate how unusual an observed diet breadth is given the branching structure of a specific host phylogeny and a particular number of host species in a diet. In order to do this, diet breadth indices can be calculated for random samples of host taxa in which the number of host taxa sampled is the same as the number utilised. The sampling should be constrained such that hosts in each sample are taken without replacement from the list of all possible hosts. Sampling proceeds until all possible diet breadth values for a particular number of host taxa have been obtained, or until a sufficiently large number of possible diet values (e.g. 10,000) has been calculated (in cases where the number of possible combinations of a given number of host taxa is especially large). Such a null-distribution approach has been implemented in a computer programme (named MSST) written by F.B. Symons and A. Rambaut (Oxford University) in the C programming language. This returns a probability for an observed diet breadth based on the null distribution for that particular number of host taxa. Where diets are examined for statistical significance in this way, it is important to correct for multiple tests, because the diet breadths for insect species which utilise the same number of hosts form part of the same null distribution of possible diet breadth values.

Combining such an approach with a comparative analysis of diet breadth would allow quantification of the extent to which a group of insects' diets are restricted compared with random diet breadth index values. For example, one might investigate whether species with smaller diet breadths are more absolutely restricted in terms of the phylogenetic diversity of their hosts than are more polyphagous species, and whether there is a threshold level of polyphagy above which diet breadths become as large as random values (or even larger than the bulk of them).

## Availability of data

Macroecological studies are often hampered by the lack of detailed phylogenies for the organisms of interest. The purpose of this paper is not to highlight or discuss this deficiency in detail but rather to present methods for analysing ideal data where these are present. Many available phylogenies possess one or more polytomies resulting from the inadequacies of phylogeny reconstruction methods. In these cases, the naive application of the methods presented here could lead to bias, because the host taxa in a polytomy remain two branch segments apart no matter how many species are involved in the multifurcation. Therefore, diet breadth indices will be progressively underestimated as the size of the polytomy increases. To avoid such bias, such nodes could either be excised from a phylogeny (where they are rare and/or do not involve many host taxa), or they could be randomly resolved and assigned branch lengths according to some model of branch length evolution. Where the latter approach is adopted, average values for diet breadths should be calculated over numerous random resolutions of the node in question. This will result in diet breadth values which are not biased by the size of polytomies. An example of the random resolution approach is provided by G.W. Beccaloni and F.B. Symons (unpublished data).

We are aware that the estimation of branch lengths is a contentious issue. Different phylogeny reconstruction algorithms and character datasets may yield different estimates of genetic distances between taxa. Where this is the case, we propose two possible solutions. A composite phylogeny estimate may be prepared with branch lengths which represent the consensus of datasets. Alternatively, analyses may be repeated using each of the different branch length estimates in turn. We feel that this latter approach is the more robust method, as any patterns which manifest themselves across all analyses are likely to reflect biological reality.

## Conclusion

The diet breadth indices suggested here perform a number of roles. Their consistent application would allow the comparison of results from different studies in a way which is currently not possible. Furthermore, by using a number of measures, it may be possible to determine what aspect of diet is responsible for an observed diet breadth (e.g. whether the number of taxa utilised or the relatedness of those taxa is the most important factor). In diet breadth studies, the distinction between studies which aim to quantify diet breadth with respect to overall plant phylogeny and those which aim to quantify diet breadth with respect to how insects distinguish host plants is fundamental.

Many of the indices discussed here have been implemented as computer programmes which facilitate their

calculation. The computer programme MSST mentioned above will be available from F.B.S. on request.

**Acknowledgements** We thank Mike Charleston, Paul Eggleton, Paul Harvey, Colleen Kelly, Ian Kitching, Gaden Robinson and two referees for valuable comments on the manuscript. F.B.S. was funded by NERC studentship GT4/94/397/L.

## References

- Becerra JX (1997) Insects on plants: macroevolutionary chemical trends in host use. *Science* 276:253–256
- Bernays EA (1998) Evolution of feeding behavior in insect herbivores. *Bioscience* 48:35–44
- Brown VK, Southwood TRE (1983) Trophic diversity, niche breadth and generation times of exopterygote insects in a secondary succession. *Oecologia* 56:220–225
- Cates RG (1981) Host plant predictability and the feeding patterns of monophagous, oligophagous, and polyphagous insect herbivores. *Oecologia* 48:319–326
- Dethier VG (1954) Evolution of feeding preferences in phytophagous insects. *Evolution* 8:33–54
- Dyer LA, Floyd T (1993) Determinants of predation on phytophagous insects: the importance of diet breadth. *Oecologia* 96:575–582
- Faith DP (1992a) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 61:1–10
- Faith DP (1992b) Systematics and conservation: on predicting the feature diversity of subsets of taxa. *Cladistics* 8:361–373
- Faith DP (1994) Phylogenetic pattern and the quantification of organismal biodiversity. *Phil Trans R Soc Lond B* 345:45–58
- Farrell BD, Mitter C (1990) Phylogenesis of insect/plant interactions: have *Phyllobrotica* leaf beetles (Chrysomelidae) and the lamiales diversified in parallel? *Evolution* 44:1389–1403
- Farris J (1979) The information content of the phylogenetic system. *Syst Zool* 28:483–519
- Fiedler K (1995) Lycaenid butterflies and plants: hostplant relationships, tropical versus temperate. *Ecotropica* 1:51–58
- Fiedler K (1998) Diet breadth and host plant diversity of tropical- vs. temperate-zone herbivores: south-east Asian and west Palaearctic butterflies as a case study. *Ecol Entomol* 23:285–297
- Futuyma DJ (1983) Evolutionary interactions among herbivorous insects and plants. In: Futuyma DJ, Slatkin M (eds) *Coevolution*. Sinauer, Sunderland, Mass, pp 207–231
- Futuyma DJ, McCafferty SS (1990) Phylogeny and the evolution of host plant associations in the leaf beetle genus *Ophraella* (Coleoptera, Chrysomelidae). *Evolution* 44:1885–1905
- Gaston KJ, Reavey D (1989) Patterns in the life histories and feeding strategies of British macrolepidoptera. *Biol J Linn Soc* 37:367–381
- Harvey PH, Pagel MD (1991) *The comparative method in evolutionary biology*. Oxford University Press, Oxford
- Hayes JL (1982) A study of the relationships of diapause phenomena and other life history characters in temperate butterflies. *Am Nat* 120:160–170
- Jaenike J (1990) Host specialization in phytophagous insects. *Annu Rev Ecol Syst* 21:243–273
- Janz N, Nylin S (1998) Butterflies and plants: a phylogenetic study. *Evolution* 52:486–502
- Joern A (1979) Feeding patterns in grasshoppers (Orthoptera: Acrididae): factors influencing diet specialization. *Oecologia* 38:325–347
- Lecointre G, Philippe H, Lè HLV, Guyader HL (1993) Species sampling has a major effect on phylogenetic inference. *Mol Phylogenet Evol* 2:205–224
- MacArthur RH (1972) *Geographical ecology: patterns in the distributions of species*. Harper Row, New York

- Mitter C, Farrell B (1991) Macroevolutionary aspects of insect-plant relationships. In: Bernays EA (ed) *Insect-plant interactions*. CRC, Boca Raton, Fla, pp 36–78
- Niemelä P, Tahvanainen J, Sorjonen J, Hokkanen T, Neuvonen S (1982) The influence of host plant growth form and phenology on the life strategies of Finnish macrolepidopterous larvae. *Oikos* 39:164–170
- Novotný V (1994) Association of polyphagy in leafhoppers (Auchenorrhyncha, Hemiptera) with unpredictable environments. *Oikos* 70:223–232
- Schoonhoven LM, Jermy T, Loon JJA van (1998) *Insect-plant biology: from physiology to evolution*. Chapman & Hall, London
- Strong DR, Lawton JH, Southwood TRE (1984) *Insects on plants: community patterns and mechanisms*. Blackwell, Oxford
- Williams PH, Gaston KJ, Humphries CJ (1994) Do conservationists and molecular biologists value differences between organisms in the same way? *Biodivers Lett* 2:67–78