Symbiosis catalyses niche expansion and diversification

Jeffrey B. Joy1,2

1Department of Biological Sciences, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6
2Centre for Interdisciplinary Research in the Mathematical and Computational Sciences (IRMACS) Burnaby, British Columbia, Canada V5A 1S6

Interactions between species are important catalysts of the evolutionary processes that generate the remarkable diversity of life. Symbioses, conspicuous and inherently interesting forms of species interaction, are pervasive throughout the tree of life. However, nearly all studies of the impact of species interactions on diversification have concentrated on competition and predation leaving unclear the importance of symbiotic interaction. Here, I show that, as predicted by evolutionary theories of symbiosis and diversification, multiple origins of a key innovation, symbiosis between gall-inducing insects and fungi, catalysed both expansion in resource use (niche expansion) and diversification. Symbiotic lineages have undergone a more than sevenfold expansion in the range of host-plant taxa they use relative to lineages without such fungal symbionts, as defined by the genetic distance between host plants. Furthermore, symbiotic gall-inducing insects are more than 17 times as diverse as their non-symbiotic relatives. These results demonstrate that the evolution of symbiotic interaction leads to niche expansion, which in turn catalyses diversification.

1. Introduction

Species interactions are a fundamental driver of evolutionary change. Competition, one form of species interaction, can drive the evolution of both novel phenotypic and species diversity [1]. Studies of other forms of species interactions have also illuminated symbiotic interactions (intimate, often long-term interaction between different species) as motivators of evolutionary change, spurring novel phenotypes, life-histories and developmental pathways [2,3]. Symbioses have played a critical role in the ecological diversification of many organisms by facilitating expansion into novel ecological niches [2,4], and this expansion is theorized to catalyse lineage diversification [2,3]. However, empirical tests explicitly linking symbioses to both niche expansion and species diversification remain scarce [5,6].

Symbiotic interactions are particularly conspicuous and important in mediating interactions between plant-feeding (phytophagous) insects and their host plants [4,7]. Symbioses between insects and fungi have evolved in a variety of taxa [8–14]. A diverse insect group known to mediate plant–insect interaction through symbiotic association with fungi is the ambrosia gall midges (Diptera: Cecidomyiidae) [9–11,15–28]. Phylogenetic analysis of the fungal symbiont places it within a monophyletic clade of Botryosphaeria dothidea fungi [29]. Botryosphaeria dothidea are cosmopolitan, typically free-living, ubiquitous plant pathogens capable of feeding on a wide array of plant families worldwide [29,30].

Fungal-symbiotic gall midge females from six genera within four morphologically well-separated taxonomic tribes transport conidia (fungal spores) and oviposit them along with their eggs into plant tissues. The gall structures induced by species of ambrosia gall midges on a variety of different plant tissue types (root, stem, bud, leaf, flower or seed) are lined internally with fungal hyphae which the developing larval gall midges use as food [10,11,15,31–35], but also for defence against natural enemies [29,30,36,37].
The genera of ambrosia gall midges display different adaptations for transportation of the fungal spores, including specialized mycangial structures or modified hairs that wrap around fungal spores [10,11,34]. As with other phytophagous insects [38–40], gall midge host-plant shifts occur most often between closely related plants relative to distantly related plants, because shifts among distantly related plants require substantially more adaptation [6,39]. Thus, gall midges are constrained in the range of possible host-plant shifts by the genetic distance between plant taxa [41,42].

Host-plant preferences are well characterized for gall midge species in 351 genera [41,43] which also may be categorized by the presence or the absence of fungal symbionts [15,30,43]. Both molecular and morphological analyses have also revealed that a small proportion (53 of the approx. 6100 + gall midge species) [43] are polyphagous (capable of using multiple host-plant species; electronic supplementary material, table S3 and figure S4). Here, using time calibrated phylogenies of gall midge host-plants, I test two hypotheses derived from evolutionary theories of symbiosis [2–4,44] and diversification [1,45] concerning the role of symbiotic interaction in niche expansion and phytophagous insect diversification: (i) plant-feeding insects engaged in symbiosis should exhibit niche expansion (display greater genetic distance between the host plants they use), when compared with non-symbiotic lineages; and (ii) such niche expansion should be concomitant with elevated diversification (figure 1).

2. Material and methods
(a) Host-plant data
Host-plant preferences for species from 351 genera of gall-inducing midges using 141 plant families were assembled and each midge taxa scored for the presence or the absence of symbiotic fungi from the literature [43]. Details of host-plant preferences and use of fungi by gall-inducing midge genera and polyphagous species is presented in electronic supplementary material, tables S1 and S2. The species richness of each gall midge genus was obtained from published sources [43] (see the electronic supplementary material, table S2).

(b) Host-plant phylogenetics
Relationships among host-plant taxa were reconstructed for gall midge genera and polyphagous gall midge species using Phylomatic [46], a tool which compiles published angiosperm phylogenies yielding a working hypothesis about their phylogenetic relationships. Where necessary host-plant phylogenies were further resolved using published phylogenies. Studies based on multiple genes were preferred and support values greater than 80 per cent were required to resolve relationships. Branch lengths of phylogenetic trees were scaled to time using the BLADJ function within Phylocom [46] and fossil calibration points from the literature [47]. The phylogenies of gall midge host-plants (see the electronic supplementary material, figures S1 and S2) are deposited in TreeBase (www.treebase.org). The maximum phylogenetic range (maximum patristic distance) of the host-plant phylogeny used by each gall midge taxa was calculated using functions in the R package APE.

(c) Statistical methods
To test the hypothesis that plant-feeding insects engaged in symbiosis should exhibit niche expansion, comparisons between fungal-symbiotic and non-fungal-symbiotic gall midges were performed at both the generic and specific taxonomic levels. Comparisons of both observed and expected numbers of polyphagous fungal-symbiotic and non-fungal-symbiotic gall-inducing midges and the observed versus expected numbers of plant families used by fungal-symbiotic and non-fungal-symbiotic gall-inducing midges were performed using $\chi^2$ goodness of fit tests. The expected numbers of gall midges were derived from
the total numbers of fungal-symbiotic and non-fungal-symbiotic gall midges and total numbers of plant families used by fungal-symbiotic and non-fungal-symbiotic gall midge species. For example, the expected number of fungal-symbiotic polyphagous gall midges is given by \((\text{total number of fungal} – \text{symbiotic gall midges/total number of gall midges}) \times \text{total number of polyphagous gall midges}\). It is conceivable that further investigation of some of the described cases of polyphagy using molecular methods will reveal some of them to be cryptic species associated with a single host plant, however, it is unlikely that these cases would be biased towards fungal-symbiotic gall midges relative to those species which do not employ fungal symbionts.

Many species of gall midges are narrowly host-specific leading to a host-phylogenetic distance of zero. An overabundance of zeros may prove problematic for traditional data modelling techniques, such as generalized linear models (GLM), these situations arise when dependent variables are comprised of more zeros than expected under a Poisson distribution [48–51]. An over abundance of zeros in dependent variables is often generated when one process gives rise to zero and another to non-zero data. In such cases, restricting-dependent variables to non-zero data may result in biased parameter estimates [48–51]. Such zero weighted datasets are well modelled with a two part zero-inflated generalized linear model termed a hurdle model [49] which accounts for the over dispersion resulting from zero-inflated data [48–51]. Host-plant phylogenetic distance (maximum patristic distance) used by non-fungal-symbiotic gall midge genera was compared with that of fungal-symbiotic gall midge genera using hurdle zero-inflated generalized linear models [48] with species richness of gall midge genera included as a model covariate as implemented in the R statistical packages pocl [52] and mgcv [53]. Host-plant phylogenetic distance used by symbiotic and non-symbiotic polyphagous gall midge species was compared via a Wilcoxon signed-rank test.

Morphologically defined taxonomic tribes [43] correspond well with a molecular phylogeny (see the electronic supplementary material, figure S3). However, the lack of a fully resolved phylogeny for Cecidomyiidae precludes the use of a traditional sister group based approach, thus comparisons of the species richness of fungal-symbiotic gall midge genera to non-fungal-symbiotic genera were performed in two ways: (i) mean species richness of fungal-symbiotic genera were compared with all genera without symbiotic fungi within their taxonomic tribes using a paired t-test and (ii) to account for remaining uncertainty in taxonomic tribes fungal-symbiotic genera were compared with all non-fungal-symbiotic plant-feeding genera in family Cecidomyiidae using a Wilcoxon signed-rank test.

That increased species richness in fungal-symbiotic lineages is a result of increased diversification rather than an effect of clade age was tested in two ways. First, the lengths of the pendant edges from the eccentric phylogeny (electronic supplementary material) of fungal-symbiotic genera were compared with those of non-fungal-symbiotic genera using the R package APE [54] and a Wilcoxon signed-rank test. Second, I developed and employed a novel diversification metric integrating niche breadth, species richness and taxon age. diversification \((D) = \text{species richness (SR) \times niche breadth (NB)/taxon age (T)} \) so, \(D = \text{SR \times NB}/T\). Thus, applied here as: \(D = \text{(gall midge genus species richness \times maximum patristic distance of plant phylogeny used by gall midge genus)/pendant edge length of gall midge genus from the gall midge phylogeny} \). \(D\) for fungal-symbiotic gall midge genera was then compared with \(D\) of non-fungal-symbiotic gall midge genera using a Wilcoxon signed-rank test. All statistical analyses were performed within the R language statistical framework [55].

First, I performed two comparisons among polyphagous gall midge species (see the electronic supplementary material, table S3) using \(\chi^2\)-tests; (i) the observed versus expected numbers of polyphagous gall midge species that do and do not use fungal symbionts; and (ii) the observed versus expected number of plant families used by polyphagous gall midge species with and without fungal symbionts. Results indicate that there are significantly more polyphagous gall midges that are symbiotic than expected \((\chi^2 = 21.62, p < 0.001)\) and that symbiotic gall midge species feed upon a greater number of plant families than expected \((\chi^2 = 38.82, p < 0.001)\). Second, the phylogenetic range of host plants used by gall midges with and without fungal symbionts was compared using both a zero-inflated generalized linear model [48] and Wilcoxon signed-rank tests. I performed this comparison both between gall midge genera with and without fungal symbionts while accounting for the species richness of gall midge genera and between known cases of polyphagous gall midge species also with and without fungal symbionts. Gall midges employing fungal symbionts to mediate interactions with host-plant species consistently use a significantly expanded range of host plants relative to non-symbiotic lineages (table 1, figures 2a,b and 3a,b).

To test hypothesis two, that fungal-symbiotic gall midge lineages are more diverse relative to lineages without such symbionts, I first compared the species richness of gall midge genera using symbiotic fungi to all related non-fungal-symbiotic lineages within their taxonomic tribes using a paired t-test. Secondly, as it is possible that a robust molecular phylogeny of family Cecidomyiidae would reveal phylogenetic rearrangements of genera among tribes the species richness of fungal-symbiotic gall midge genera was compared with all other plant-feeding genera within the family. Consistent with prediction (2), gall midge lineages using symbiotic fungi to mediate insect–plant interactions are significantly more diverse than both their tribal relatives lacking fungal symbionts \((t = 21.71, d.f. = 3, p < 0.0005; \text{figure } 3c\) electronic supplementary material, figure S4) and all plant-feeding genera within the family \((W = 97.5, p < 0.0001)\). As judged by the available sequence data and a Wilcoxon test fungal-symbiotic gall midge genera are not significantly older when compared with their available non-fungal-symbiotic relatives \((W = 76.3, p = 0.75)\). Comparison of diversification \((D)\) of fungal-symbiotic gall midge genera with \(D\) of non-fungal-symbiotic gall midge genera suggests that symbiotic genera are diversifying faster \((W = 132, p = 0.0008; \text{figure 4})\).

4. Discussion

In accord with both predictions these results consistently support a role for symbiosis in niche expansion and diversification of phytophagous insects. Phytophagous insects which have evolved symbiotic association with fungi exhibit niche expansion, in the form of use of a broader range of host-plant taxa, because their fungal symbionts are cosmopolitan plant pathogens capable of digesting dozens of plant families worldwide [30,33]. Buffering interactions with plant taxa through symbiotic interaction with generalist fungi may thus relax phylogenetic constraints on the host-plant preferences of phytophagous insects imposed by characteristics of their host-plant lineages such as host-plant chemistry [57], natural enemies [36,58] or plant defences [59].

3. Results

I tested hypothesis (i) that plant-feeding insects with symbiotic associations should display niche expansion in two ways.
Key innovations facilitate the rapid migration from one peak in the adaptive landscape [60,61] to another perhaps higher peak [62] catalysing diversification [63]. For plant-feeding insects, such as gall-inducing midges, higher peaks in the adaptive landscape constitute those with more ecological opportunity in the form of more available niches (unexploited host-plant taxa, plant parts or time periods of plant growth) [41]. Fewer competitors or fewer natural enemies [36]. Incorporation of fungal symbionts into the life cycle of plant-feeding insects is thus a key innovation allowing transitions across a deep valley in the adaptive landscape, a shift to a genetically distant host-plant taxon, to be similarly tractable to shifts between genetically similar host-plant taxa. Thus, by facilitating the colonization of distantly related host-plant taxa, symbioses may allow insect lineages to simultaneously escape the limitations present at smaller scales, such as local competition for limited niches, and take advantage of the ecological opportunities that accompany shifts to novel adaptive zones [64], catalysing diversification at larger scales. Colonization of such novel adaptive zones is likely accompanied by the opportunity to diversify sensu Simpson [65] ‘...more or less simultaneous divergence of numerous lines all from much the same ancestral type into different, also diverging adaptive zones’, [65, p. 223], predicting elevated diversification of fungal-symbiotic insect lineages relative to lineages without fungal associations. Thus, among phytophagous insects, colonization of disparate adaptive zones may facilitate both phenotypic divergence [40] and diversification through division of available niche space such as closely related host-plant species, novel plant parts and time periods [36,41].

The 548 fungal-symbiotic gall midge species with well-described host-plant preferences are roughly evenly divided between tropical (190 species), temperate (215 species) and arid (143 species) biogeographical regions. Furthermore, only 24 fungal-symbiotic gall midge species use plant families that are not also used by non-fungal-symbiotic gall midge species suggesting the effect of symbiosis on niche

---

**Figure 2.** Plot of zero-inflated (hurdle) generalized linear model comparing phylogenetic range used by symbiotic (orange dashed lines) and non-symbiotic (green solid lines) gall midge genera, in which the response variable is phylogenetic breadth of the host-plant phylogeny used (host-plant phylogenetic distance \(\sim \log(\text{gall midge species richness})\)) + symbiotic status (symbiotic status). For symbiotic and non-symbiotic gall midge genera of similar species richness the symbiotic genera consistently use a greater proportion of the host-plant phylogeny. (a) Fitted values for binomial portion of hurdle model. (b) Fitted values for Poisson portion of hurdle model. In both (a) and (b) CIs are plus or minus twice the s.e. from model fit. (Online version in colour.)

**Table 1.** (a) Results of zero-inflated generalized linear model (hurdle model) comparing phylogenetic range (maximum patristic distance) of host plants (HP) used by symbiotic and non-symbiotic gall midge genera. The sample size \(n\), \(\beta\), \(\beta\) s.e., \(z\)-score and \(p\)-value are provided for both the Poisson and binomial parts of the model for each independent variable. (b) Wilcoxon comparison of the breadth of the host plant phylogeny (maximum patristic distance) used by polyphagous gall midge species using a fungal symbiont with the breadth of the phylogeny used by polyphagous gall midge species without fungal symbionts.

<table>
<thead>
<tr>
<th>model and response variables</th>
<th>independent variables</th>
<th>(n)</th>
<th>(\beta)</th>
<th>(\beta) s.e.</th>
<th>(z)</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) gall midge genera: (i) Poisson part</td>
<td>HP phylogenetic range (intercept)</td>
<td>350</td>
<td>6.17516</td>
<td>0.009</td>
<td>654.416</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>HP phylogenetic range symbiont</td>
<td>350</td>
<td>0.188</td>
<td>0.019</td>
<td>9.652</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>HP phylogenetic range gall midge genera species richness</td>
<td>350</td>
<td>0.059</td>
<td>0.003</td>
<td>15.886</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(a) gall midge genera: (ii) binomial part</td>
<td>HP phylogenetic range (intercept)</td>
<td>350</td>
<td>1.645</td>
<td>0.146</td>
<td>11.267</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>HP phylogenetic range symbiont</td>
<td>350</td>
<td>3.254</td>
<td>1.105</td>
<td>2.944</td>
<td>0.003</td>
</tr>
<tr>
<td>(b) polyphagous gall midge species</td>
<td>HP phylogenetic range non-symbiotic</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HP phylogenetic range symbiont</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
expansion and diversification is not simply a correlate of particular host-plant families or biogeographical regions. While good taxonomic characters delineate the fungal-symbiotic genera, different fungal-symbiotic genera have unique adaptations associated with symbiosis, and the preliminary molecular phylogeny matches well with taxonomy it is impossible to say for certain that the genera are monophyletic without a completely resolved molecular phylogeny. Furthermore, comparison of the richness of genera is not ideal nor is the estimation of lineage age without a complete phylogeny. Thus, a complete molecular phylogeny for the group would provide a more robust framework for testing the hypotheses here thereby substantially strengthening the inference that symbiosis catalysed diversification of this group.

Taken together these results support predictions derived from evolutionary theories of symbiosis and diversification that symbiotic interaction catalyses niche expansion and diversification. The importance of symbiotic interaction in diversification is unlikely to be exclusive to gall midge–fungal associations or to plant-feeding insects. For instance, it is likely that microbial mutualists also promote diversification through expansion of ecological opportunities available to their hosts in other taxa [4,6]. Thus, studies of the consequences of symbiosis for diversification in other taxonomic groups and other contexts of symbiosis would further illuminate the role played by symbiosis in diversification.

I am grateful to Bernard Crespi, Laura Weir, Patrik Nosil, Christine Parent, Arne Mooers and Will Stein for comments that greatly improved this manuscript. I thank Ruth Joy for helpful statistical advice and Suzanne Kuhnholz for translation of German manuscripts into English. I am greatly indebted to the SFU FAB* lab for helpful discussion. This work was supported by an NSERC discovery grant to Arne Mooers and a CIHR operating grant to Jamie Scott. The Interdisciplinary Research in Mathematics and Computing Sciences Centre at Simon Fraser University provided facilities.

References


