Testing host-plant driven speciation in phytophagous insects: a phylogenetic perspective

Emmanuelle Jousselin* ¹, Marianne Elias ²

¹ CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, Univ Montpellier, Montpellier, France
² Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 57 rue Cuvier, CP 50, 75005 Paris, France.

*corresponding author: emmanuelle.jousselin@inra.fr

Keyword: coevolution, herbivory; host-plant specialization, phylogeny, speciation.

Abstract

During the last two decades, ecological speciation has been a major research theme in evolutionary biology. Ecological speciation occurs when reproductive isolation between populations evolves as a result of niche differentiation. Phytophagous insects represent model systems for the study of this evolutionary process. The host-plants on which these insects feed and often spend parts of their life cycle constitute ideal agents of divergent selection for these organisms. Adaptation to feeding on different host-plant species can potentially lead to ecological specialization of populations and subsequent speciation. This process is thought to have given birth to the astonishing diversity of phytophagous insects and is often put forward in macroevolutionary scenarios of insect diversification. Consequently, numerous phylogenetic studies on phytophagous insects have aimed at testing whether speciation driven by host-plant adaptation is the main pathway for the diversification of the groups under investigation. The increasing availability of comprehensive and well-resolved phylogenies and the recent developments in phylogenetic comparative methods are offering an unprecedented opportunity to test hypotheses on insect diversification at a macroevolutionary scale, in a robust phylogenetic framework. Our purpose here is to review the contribution of phylogenetic analyses to investigate the importance of plant-mediated speciation in the diversification of phytophagous insects and to present suggestions for future developments in this field.
Introduction

The idea according to which new species arise through adaptation to different ecological niches constitutes the core of Darwin’s work. This process is now termed ecological speciation and its study has become an intense field of research in evolutionary biology (Nosil, Crespi, Sandoval 2002; Rundle, Nosil 2005; Schluter 2009; Nosil 2012). **Phytophagous insects** have always been at the forefront of these investigations (Drès, Mallet 2002; Funk, Filchak, Feder 2002; Matsubayashi, Ohshima, Nosil 2010; Elias et al. 2012; Forbes et al. 2017). The hypothesis of ecological speciation resulting from divergent selection exerted by host-plants was put forward a long time ago to explain the formation of new species of insects (Brues 1924). There are several model systems on which this scenario has been explored. One text-book example of host-plant driven incipient speciation is the apple maggot (*Rhagoletis pomonella* complex) in which the evolution of new feeding preferences on the recently introduced domesticated apple (*Malus pumila*) has supposedly led to the emergence of specialized host races (Bush 1975; Berlocher 2000; Powell et al. 2014), Stick insects, leaf beetles (Nosil et al. 2012); (Rundle et al. 2000), butterflies (McBride, Singer 2010), and the pea aphid also star among model systems in the study of host-driven speciation (Via, Bouck, Skillman 2000); (Caillaud, Via 2000); (Peccoud et al. 2009);(Smadja et al. 2012). In all these examples, the speciation scenario hypothesizes that: 1) the restricted utilization of distinct sets of host-plant species by insect populations is the result of adaptive trade-offs; 2) hybrids with intermediate phenotypes (in terms of traits involved in host-plant adaptation*) fare poorly on parental host-plants and are selected against and therefore gene flow between populations is reduced; 3) gene flow can further be reduced through the evolution of assortative mating, especially when host-plants also represent mating sites. In support of this scenario, many studies show the existence of genetically differentiated host races in insect species. Some studies have quantified selection against hybrids (McBride, Singer 2010; Gow, Peichel, Taylor 2007) and some studies have uncovered genomic regions that determine host-plant preference and performance on alternative hosts (Egan, Nosil, Funk 2008; Smadja et al. 2012).

The role of host-plant-mediated speciation in the diversification of phytophagous insect lineages is also largely emphasized in the literature on large-scale patterns of insect diversity: macroevolutionary perspectives on phytophagous insect evolution have attributed their extraordinary diversification to selective responses to their host-plants (Ehrlich, Raven 1964; 1 See glossary
Yokoyama 1995; Winkler, Mitter 2008; Janz 2011). However, these macroevolutionary scenarios are often presented in the literature as narratives for specific lineages commenting a phylogenetic reconstruction of the history of host-plant associations. Many phylogenetic studies still fail to clearly formulate hypotheses and predictions about the speciation processes that underlie the observed patterns and the role played by host-plant adaptation in those. The reason might be that the macroevolutionary patterns that arise when host-plant specialization* is the driver of speciation events are not always clear. There is no review on what to expect and how to formally test these predictions.

The increasing availability of robust molecular phylogenies and recent developments in phylogenetic comparative methods are offering an unprecedented opportunity to test evolutionary hypotheses in a robust phylogenetic framework. Our purpose here is to present the macroevolutionary scenarios for the diversification of phytophagous insects that have been put forward in the literature, decipher the role that ecological speciation driven by host-plant adaptation play in them and synthetize predictions from these scenarios. We then identify tools from the “comparative phylogenetic toolbox” that provide ways to test some of these predictions. This toolbox can be divided into three compartments:

1) comparisons of the phylogenetic histories of insects and their associated plants: the congruence (in terms of dates of divergence and branching patterns—and) of the phylogenetic histories of plant-feeding insects and their host-plants can be tested in robust statistical frameworks and illuminate how herbivores track the diversification of their hosts;

2) ancestral character state reconstructions: the evolutionary trajectory of host-associations, host breadth and host-plant adapted traits can be inferred using ancestral character state reconstruction methods and statistical tests can determine whether their distribution throughout the phylogenetic trees follow the predictions of scenarios involving ecological speciation mediated by host-plant adaptation;

3) diversification analyses: the recent developments of methods to study the diversification dynamics* of entire clades using phylogenetic trees provide ways to test how shifts to new host-plant species or changes in host breadth have impacted speciation/diversification rates in phytophagous insects.

We review papers that have adopted these approaches. We then present suggestions for future research that should help linking microevolutionary studies on host-plant adaptation and macroevolutionary perspectives on phytophagous insect diversification.

I Macroevolutionary scenarios of phytophagous insect diversification
I.1) Escape and radiate (Figure 1 a)

More than 50 years ago, Ehrlich & Raven (Ehrlich, Raven 1964) put forward a macroevolutionary scenario that inspired most of the current research on plant-feeding insect diversification: it is known as “Escape and radiate” (Thompson 1989). They hypothesized that when insects acquire the ability to circumvent the chemical defenses of a plant group, it promotes their rapid diversification by ecological release, i.e. the availability of novel resources and reduction in direct competition. Insects undergo an adaptive radiation*. In this scenario, adaptation towards host-plants is the driving force of insect species formation. The “Escape and radiate” scenario also hypothesizes that, in response to phytophagous insect predation, plants acquire novel chemical defenses which allow them in turn to diversify very rapidly (Marquis et al. 2016). Ehrlich & Raven’s seminal study suffers from several shortcomings that have been pinpointed before (Janz 2011). First, although the authors frame their theory within the concept of adaptive radiation, they do not explicitly lay out some speciation mechanisms for both partners of the interaction. Following their scenario, a trade-off in resource use and specialization towards specific host-plants is necessary to explain the formation of numerous insect species (i.e. species radiation) following the capture of a new host-plant lineage. Such a trade-off is not mentioned in the original paper (Janz 2011). In addition, as underlined by contemporary researchers of Ehrlich and Raven’s, it is difficult to conceive how the selection pressures exerted by insects on plant defences can drive plant speciation (Jermy 1976), (Jermy 1984). Plant traits that reduce phytophagous insect attacks are rarely linked with reproductive isolation between plant populations (but see (Marquis et al. 2016) for a review of scenarios of herbivore-induced speciation in plants) and the evidence for bursts of speciation in plants following the evolution of chemical defence is scant (Futuyma, Agrawal 2009). However this study has been and remains a great source of inspiration for studies on the diversification of plant-insect associations. This is probably because it is one of the first studies that attempts to explain how microevolutionary processes (host-plant adaptation) translate into macroevolutionary patterns (radiation onto newly acquired plant lineages). Several predictions that can be tested on phylogenetic trees arise have been derived from the Escape & Radiate scenario (Table 1).

In the years following its publication, “Escape and radiate” was often interpreted as generating cospeciation* patterns; however it is now recognized that it rather predicts the sequential speciation of insects onto an already diversified plant lineage (Janz 2011; Suchan, Alvarez 2015). According to this prediction: 1) the reconstruction of the history of host-plant associations on the phylogenetic trees of insects should reveal host-plant conservatism*, i.e. the use of –related plant species by related insects (Winkler, Mitter 2008); 2) the phylogenies of
herbivorous insects and their host-plants should be more congruent than expected by chance and the diversification of the insects should lag behind that of their host-plants; this is called sequential evolution (Jermy 1984) or “host tracking”. Nevertheless, when the association between insects and their host plants is species-specific, a pattern of cospeciation can be expected through simple co-vicariance: geographic barriers affect the differentiation of populations of interacting lineages in a similar way and cause simultaneous speciation events (Althoff et al. 2012; Brookes et al. 2015; Martínez-Aquino 2016). In these cases, it is geographic isolation and not natural selection that initiates the reproductive isolation of insect populations and subsequent speciation. However, the specificity of the insects and host-adapted traits enhance the probability of shared vicariant events.

The diversification dynamics of insects should follow the typical pattern of adaptive radiations (Janz 2011), i.e. they should show an acceleration of speciation rate upon the capture of new plant lineages or the evolution of detoxification mechanisms (Wheat et al. 2007) and then slow down when their niches are saturated (when species diversity is reaching the carrying capacity of the host-plant lineage) (Rabosky, Lovette 2008). Furthermore, the capture of a species-rich clade of plants should result in higher speciation rates than the capture of lineages encompassing less species (Roskam 1985).

1.2) The ‘Oscillation Hypothesis’ (Figure 1b)

The “Escape and radiate” scenario was revisited more than a decade ago by Janz and collaborators (Janz, Nylin, Wahlberg 2006; Janz, Nylin 2008; Nylin, Janz 2009). Using butterflies as study systems, they stated that expansions in diet breadth followed by specialisation onto new host-plant species constantly fuel the diversification of phytophagous insects. This has been termed the ‘Oscillation Hypothesis’ (Janz & Nylin 2008). It stipulates that transitions towards a generalist diet generally open up a new adaptive zone, which favours the capture of new host-plants. In this scenario, expansions in diet are enabled by the phenotypic plasticity of insects with respect to host-plants (Nylin, Janz 2009). Population fragmentation and their specialisation onto newly captured host-plants then lead to the formation of new species. Hence this scenario explicitly predicts that species formation results from insect populations evolving towards the utilization of a restricted set of host-plants. Nevertheless, it suggests that this specialization process is often a consequence of the geographic isolation of generalist insect populations in areas inhabited by different host-plant species (Janz, Nylin, Wahlberg 2006). Therefore the “Oscillation hypothesis” does not necessarily postulate that natural selection is the main driving force of species formation.
**Escape and Radiate** (Erlich & Raven 1964)

Host shift \[\rightarrow\] Radiation on related plants

Evolution of detoxification

**Oscillation hypothesis** (Janz & Nylin 2008)

Broadening of feeding diet \[\rightarrow\] Specialization on sub-sets of host plants
Figure 1: Schematic illustrations of the macroevolutionary scenarios of phytophagous insect diversification.

Musical chairs (Hardy & Otto 2014)
Speciation through successive host switches
However, subsequent papers quoting this scenario emphasize the central role of host-plant specialisation (Hardy, Otto 2014; Nakadai 2017; Wang et al. 2017).

Most of the predictions from the *Escape and Radiate* scenario are valid under the "Oscillation hypothesis." However, the latter hypothesis yields several new predictions (Table 1):

- **generalist** diets should be “transient and repeatedly disappear in favour of specialization onto a limited set of related plants” (Nylin, Slove, Janz 2014);
- gains of new host plants are associated with host breadth expansion (Janz et al. 2001);
- the amplitude of the oscillation determines the number of potential host-plant species, therefore, insect clades with the most diverse host-use (the highest number of host-plant species) are expected to be more speciose than clades using *less-fewer* host species (Janz, Nylin, Wahlberg 2006); and along the same lines insect clades that encompass species that exhibit large host breadths should have higher *diversification speciation* rates (Weingartner, Wahlberg, Nylin 2006; Hardy, Otto 2014);
- shifts from a generalist diet to a specialist one should be associated with an acceleration of *diversification speciation* rates. In other words, patterns of diversification should follow a model where cladogenetic events are associated with host breadth reduction (Hardy, Otto 2014);
- generalists have larger geographic ranges as they are able to colonize more habitats and can expand more easily (Slove, Janz 2011).

**I.3) The Musical Chairs (Figure 1c)**

As opposed to the *Oscillation hypothesis*, Hardy and Otto (2014) have proposed an alternative scenario in which recently put forward a scenario where speciation in herbivorous insects is driven by host-switching rather than transitions in diet breadth by specialisation onto newly captured host plants without any transitions in diet breadth: insects speciate by successive switches to new host plants. The authors named their scenario “the *Musical Chairs*” (Hardy, Otto 2014). In this scenario, phytophagous insects speciate by the sequential capture of new host-plants and therefore the speciation of a plant-feeding insect lineage is merely driven by the number of plants it can colonize. Host-driven speciation remains at the centre of this hypothesis. They explicitly state that specialization onto a few host plant species explains species diversification in phytophagous insect.

The “*musical chairs*” hypothesis yields several predictions that differentiate it from the “*Oscillation hypothesis*” previous macroevolutionary scenarios (Table 1):
contrary to the *Escape* and *Radiate* hypothesis, the capture of a new host plant lineage does not initiate a radiation; speciation rates should be higher in lineages showing no conservatism in host-plant associations (Hardy & Otto 2014); contrary to the oscillation hypothesis, gains of new hosts are not associated with host breadth expansion (Hardy 2017); host breadth contraction is not associated with cladogenetic events (Hardy, Otto 2014) and overall there should be a negative correlation between host-plant breadth and speciation rates (Hardy, Otto 2014); host breadth contraction is not associated with cladogenetic events (Hardy & Otto 2014); speciation rates are higher in insect lineages that exhibit lability* in host-plant associations; lineages showing many transitions from generalist to specialist feeding diets should not be more speciose than lineages that only encompass specialist species (Hardy & Otto 2014).

The lack of connection between host breadth contraction and speciation events and the negative association between host-breadth and speciation rates clearly differentiate the *Musical chairs from the Oscillation hypothesis*. However, several authors have pointed out that it will be difficult to tell apart the “*Musical Chairs*” from “the *Oscillation Hypothesis*”. If generalism generalist feeding diets are indeed ephemeral as expected when specialization towards host-plants is adaptive, it will be difficult to reconstruct its history accurately on phylogenetic trees (Hardy, Peterson, Normark 2016; Janz et al. 2016). Consequently, the relationships between host-range size and speciation rates will be difficult to explore and the set of predictions that differentiate the *musical chairs from the oscillation hypothesis* will not always be testable.

Phylogenetic comparative methods (Pennell, Harmon 2013) including increasingly sophisticated diversification models (Rabosky 2006; Stadler 2013; Stadler, Bokma 2013; Rabosky, Goldberg 2015; Beaulieu, O’Meara 2016; O’Meara, Beaulieu 2016) can now be deployed to reconstruct ancestral character states, investigate the diversification dynamics of insect lineages and test whether shifts in diversification rates are associated with transitions in character states. Below we review how these methods have been used to investigate the evolution of plant/insects associations and test the predictions of host-driven speciation scenarios.
Table 1: Predictions from different scenarios involving host-plant driven speciation scenarios: the first column indicates the evolutionary hypotheses tested and the headers of the other columns indicate phylogenetic comparative approaches used for testing them.

<table>
<thead>
<tr>
<th>Macroevolutionary scenarios</th>
<th>Insect and host tree comparison</th>
<th>Evolution of host associations</th>
<th>Evolution of host breadth</th>
<th>Insect diversification dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speciation driven by host-plant specialization</td>
<td>Cospeciation if the insects play a role in their host-plant reproductive isolation</td>
<td>No overlap in host use among sister-species (Nyman et al. 2010).</td>
<td>Predominance of specialists over generalists (Janz et al. 2001; Winkler &amp; Mitter 2008).</td>
<td>- herbivorous insects are more diverse than their non-herbivorous relatives (Mitter et al. 1988) - the diversification dynamic of phytophagous insects follows a pattern of adaptive radiation (Janz 2011) - the number of species within an insect clade positively correlates with the number of host-plant species (Janz 2006)</td>
</tr>
<tr>
<td>Escape and radiate</td>
<td>Host tracking (Jermy 1976, Mitter &amp; Brooks 1983).</td>
<td>Phylogenetic conservatism of host-plant lineages or host-plants with similar defences (Winkler &amp; Mitter 2008).</td>
<td>No prediction</td>
<td>Increase in speciation rates upon the capture of new plant lineages or detoxification mechanisms (Wheat et al. 2007; Fordyce 2010).</td>
</tr>
<tr>
<td>Oscillation</td>
<td>Potentially host tracking following the capture of a new host-plant lineage.</td>
<td>Conservatism of host-plants following the capture of a new plant lineage.</td>
<td>- recurrent transitions in lability in host breadth (Janz et al. 2001, Janz &amp; Nylin 2008) - gains of new host lineages are preceded by host breadth expansion (Janz et al. 2001) - host breadth larger in species with large geographic range - positive correlation between diet breadth and geographic range (Slove &amp; Janz 2011) - speciation events associated with shifts from generalist to specialist (Hardy, Otto 2014)</td>
<td>- clades including generalist species are more speciose than clades with only specialists (Weingartner, Wahlberg, Nylin 2006) - speciation rates higher in lineages showing high lability in host breadth (Hardy, Otto 2014)</td>
</tr>
</tbody>
</table>
| Musical chairs | No prediction | Less conservatism of host plant - High lability in more speciose lineages breadth. (Hardy, Otto) | Few transitions in host breadth is not labile. (Hardy & Otto 2014; Hardy 2017) | - no positive association between speciation rates and host breadth lability (Hardy, Otto 2014) - speciation rates positively correlated
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2014)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>with lability in host associations switching rates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Hardy, Otto 2014)</td>
</tr>
</tbody>
</table>
II Phylogenetic approaches for testing ecological speciation scenarios

II.1) Comparing the phylogenies of plants and insects

Many **Several** phylogenetic studies have compared the phylogenies of herbivorous insect and their host-plants and **Some** have investigated cospeciation and host tracking patterns. They using have used **dedicated** tools such as tree reconciliation analyses (Page 1994; Conow et al. 2010) and distance-based methods for tree comparisons (Legendre, Desdevises, Bazin 2002). These tools statistically test the null hypothesis that the phylogenies of insects and their associated plants are more congruent than expected by chance and that speciation events are simultaneous. As stated above, this is only expected in species specific plant/insect interactions. A pattern of cospeciation As predicted by early taxonomic studies (Ramirez 1974; Wiebes 1979), cospeciation has been evidenced in figs and their phytophagous pollinating wasps found between figs and their phytophagous pollinating wasps but the degree of phylogenetic congruency observed varies according to taxonomic groups (see review by de Vienne, 2013 on cospeciation studies and —Cruaud, 2012) as predicted by early taxonomic studies on this biological system (Ramirez 1974; Wiebes 1979), but also between figs and some of the non-pollinating galling wasps that are highly specific to their hosts (Jousselin et al. 2008). Reciprocal adaptations of plants and insects (i.e. coevolution*) have been unravelled in this study system (Jousselin, Rasplus, Kjellberg 2003; Weiblen 2004). However, it is not known whether the cospeciation patterns observed are the sole result of this coevolution (i.e. whether reciprocal selection exerted by both partners) has driven the reproductive isolation of interacting populations (Althoff, Segraves, Johnson 2014; Hembry, Yoder, Goodman 2014) or whether matching speciation events have arisen through co-vicariance. In the other iconic model system for the study of plant/insect coevolutionary diversification*, the Yucca–Yucca moth interaction (Pellmyr 2003), pollinating and non-pollinating the moths phylogenetic histories parallel some parts of the evolutionary history of their hosts. Some studies suggest that these patterns are the results of coevolution (Godsoe et al. 2009) while others hint towards co-vicariance (Althoff et al. 2012; Kawakita 2004 #1058).

In both interactions, the fact that the phytophagous insects are specific pollinators of their host-plants and breed inside them necessarily links the reproductive success of the two partners and increases the likelihood of concomitant speciation events. Therefore host-plant adapted traits are certainly pivotal in the speciation process of these insects.

**Studies** have unravelled a pattern of investigated **phylogenetic tracking** between phytophagous insects and their host-plants (e.g. suggesting
that phytophagous insects in different orders (Coleoptera, Lepidoptera, Hemiptera) have speciated by switching and specialising onto different subsets of a newly captured plant lineage, have shown that insects rarely partly mimicking their host-plant phylogenies. All these studies give credit to the Escape and Radiate scenario. (see review by de Vienne, 2013, Winkler, 2008 #; a third of the 20 studies reported in these two reviews found partial patterns of congruence between insects and hosts plant phylogenies) (Suchan, 2015 #3814). In many plant/insect interactions, a simple observation of the patterns of distribution of host plants in the insect phylogeny can actually rule out cospeciation or phylogenetic tracking and studies on plant-insect interactions have moved away from cospeciation studies such as.

The authors have then simply compared the timing of divergence of plants and associated insects (e.g. Gómez-Zurita, 2007; Leppanen, 2012; Brandle et al. 2005; López-Vaamonde et al. 2006; Pena, Wahlberg 2008; McKenna et al. 2009; Stone et al. 2009; Kergoat et al. 2011; Segar et al. 2012; McLeish, Miller, Mound 2013; Wahlberg, Wheat, Pena 2013; Kergoat et al. 2015; Vea, Grimaldi 2016). Most of these studies suggest delayed (but sometimes rapid) colonization of already diversified groups of plants by insects groups at different temporal scales. They are generally based on mere qualitative comparisons of dates of divergence obtained from fossil calibrated phylogenies of both plants and insects, but can also include thorough statistical comparisons of dates obtained through phylogenetic methods (Loss-Oliveira, Aguiar, Schrago 2012; McLeish, Miller, Mound 2013). They are generally framed as supporting the Escape and radiate theory. However these studies do not give any information on the speciation process behind the diversification of the insect lineages studied; they merely indicate the timing of host plant colonization. The comparison of the diversification dynamics of both herbivorous insects and their host-plants provide a more direct test of host-driven speciation hypotheses: under host-driven adaptive radiation insects diversification dynamic is expected to roughly follow the diversification of its host—plant lineage. These can be investigated through diversification analyses (Kergoat et al. 2018) (see II. 3.2), but can also include the comparison of the fossil records of both insects and plants (Labandeira, 1994; Labandeira, 2013). In general, the studies of fossil assemblage are
decoupled from phylogenetic studies of plant-/insect associations. We advocate for more combining analyses: fossil and phylogenetic evidence.

II. 2) Reconstructing the evolutionary trajectory of traits involved in host-plant use

II.2.1) Evolution of host associations

Phylogenetic inferences are now widely used to reconstruct the evolutionary trajectories of phenotypic traits throughout the diversification of a lineage. Most phylogenetic studies of phytophagous insects map the history of host association onto the resulting trees— at different taxonomic levels (i.e., host plant order, family, genus depending on the level of host specialization of the insect clade studied, multistate characters are used when species are polyphagous). These reconstructions often generally show some level of host conservatism* (see Winkler & Mitter 2008 for a review). These assertions stem from mere observations of the reconstructions but numerous studies now include statistical tests. These include the permutation tail probability test (PTP, Faith, Cranston 1991 e.g., Kelley and Farrell 1998), or some index of phylogenetic signal such as the lambda ($\lambda$) of (Pagel 1999) lambda ($\lambda$) (see (Winkler, Mitter 2008) for a review and for instance Leppanen et al. 2012; Wilson et al. 2012) for statistical evidence of host conservatism on respectively: sawflies, leaf-mining moths and geometrid moths for more recent studies).

Host conservatism is often interpreted as following the predictions of “Escape and Radiate” and therefore evidence that speciation was promoted by host-plant specialization. However, showing that related insects feed on related plants does not say much about the process that has generated this pattern nor connects mechanistically host-plant use evolution to speciation. The use of vague wording such as host-associations favour or constrain speciation is commonly found when discussing host-conservatism in the literature and it is difficult to actually conclude from these studies that specialization towards one or a few plants species is the main pathway towards the formation of new phytophagous insect species. The pattern of “host conservatism” is in agreement with a scenario in which insects have radiated onto a plant lineage but it could also suggest that host-plant shifts are not important promoters of speciation events.

A more direct estimation of the contribution of host-plant adaptation in the speciation process consists in inferring the frequency of host-plant shifts in relation to speciation events. If adaptation to different ranges of host-plants drives reproductive isolation and speciation, it follows that insect sister species* should partition host-plant resources: i.e. they should show no or little overlap in the plant species they use (Jermy).
To investigate factors driving speciation, early studies have conducted in a genus of bark beetles (Aphanarthrum), Jordal & Hewitt (2004) simply compared host use of sister species comparisons of host ranges. More recent studies have and estimated that only two out of twelve cladogenetic events could be associated with host shifts. Nyman et al. (2010) used a phylogeny of sawflies belonging to the Nematinae (Hymenoptera) and reconstructed the evolution of their insects ecological niches (defined as the combination of feeding habits and host-plant families species or families). They showed that the number of niche shifts represented at the most 60% of the number of associated with speciation events (Table 2.1). Comparisons of alternative models of evolution of host-use have also been conducted. They also demonstrated that the probability that sister species overlapped in their niche decreased with time since the speciation event, suggesting that more recently diverged species have more chance of sharing host-plant species and thus have probably not differentiated via host shifts. Jousselin et al. (2013) adopted a similar approach on a genus of conifer-feeding aphids (Hemiptera). Species in this genus generally feed on one or few species and ecological niches were defined as a combination of “plant organ x host-plant species”. They—All studies but one showed that the numbers of niche shifts observed only generally represented less than 2050% of the speciation events and were generally lower than expected if the niches were randomized onto the phylogeny. This suggests that ecological speciation is not the main process behind the diversification of these lineages.

Authors have also observed that host-use differentiation occurred at the root of the trees and therefore concluded that it played a minor role in recent speciation events, suggesting that ecological speciation was not the main process behind the diversification of this aphid lineage. Recently Winkler et al. (2018) showed that species splits associated with niche differentiation (host-plant species and/or type of plant tissue attacked) were less numerous than geographic shifts throughout the phylogeny of a genus of tropical fruit-flies (Blepharoneura). To quantity niche overlap at cladogenetic events, Nakadai & Kawatika (2016) used an approach that resemble Disparity Through Time studies (Harmon et al. 2003). They computed a dissimilarity index that calculates the difference in host use across the different nodes of the phylogeny and tested whether closely related species share a more similar range of host-plant than expected by chance. They show that changes in host-plant use are concentrated at the root of the tree and play a minor role in recent speciation events.
Table 2: Summary of studies testing the predictions of host-driven speciation scenarios using phylogenetic methods. Cospeciation studies and studies on host-plant conservatism are not included as they are already synthetized in respectively: de Vienne et al. 2013 and Winkler & Mitter 2008.

<table>
<thead>
<tr>
<th>Predictions tested</th>
<th>Taxa</th>
<th>Reference</th>
<th>Approaches</th>
<th>Results</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Partitioning of host plants at</td>
<td>Aphanartum (25 spp.)</td>
<td>Jordal &amp; M Hewitt 2004</td>
<td>Sister species comparison of host ranges</td>
<td>Ø</td>
<td>Geographic isolation is more important than host switching in speciation events</td>
</tr>
<tr>
<td>speciation events</td>
<td>Nematinae (125 spp.)</td>
<td>Nyman et al. 2010</td>
<td></td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blepharoneura (49 spp.)</td>
<td>Winkler et al. 2018</td>
<td>Investigate niche shifts at speciation nodes through ancestral state</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cinara (76 spp.)</td>
<td>Jousselin et al. 2013</td>
<td>reconstruction</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caloptilia (13 spp.)</td>
<td>Nakadai &amp; Kawakita 2016</td>
<td>Niche dissimilarity index through time</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neodiprion (19 spp.)</td>
<td>Linnen &amp; Farrell 2010</td>
<td>Comparisons of models of evolution of host use (speciational model vs</td>
<td>√</td>
<td>Host shift (in allopatry) induces speciation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>gradual)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Host breadth is labile</td>
<td>Dendroctonus (19 spp.)</td>
<td>Kelley &amp; Farrell 1998</td>
<td>Phylogenetic signal of host breadth</td>
<td>√</td>
<td>Repeated broadening of host breadth</td>
</tr>
<tr>
<td></td>
<td>Lymantrinae (55 genera)</td>
<td>Wang et al. 2017</td>
<td>Phylogenetic signal of host breadth and transition rates estimation</td>
<td>√</td>
<td>oscillation</td>
</tr>
<tr>
<td></td>
<td>Nymphalidae (551 spp.)</td>
<td>Nylin et al. 2014</td>
<td>Phylogenetic signal of host breadth and transition rates estimation</td>
<td>√</td>
<td>Specialization is not a dead-end</td>
</tr>
<tr>
<td></td>
<td>Nymphalini (31 spp.)</td>
<td>Janz et al. 2001</td>
<td>Visualization of reconstruction of host breadth</td>
<td>√</td>
<td>Repeated broadening of host breadth, oscillation</td>
</tr>
<tr>
<td></td>
<td>15 insect groups</td>
<td>Nosil &amp; Mooers 2005</td>
<td>Estimation of transition rates in host breadth</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Boloria (37 spp.)</td>
<td>Simonsen et al. 2010</td>
<td>Visualization of reconstruction of host breadth</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nymphalidae</td>
<td>Hamm &amp; Fordyce 2016</td>
<td>Phylogenetic signal of host breadth</td>
<td>Ø</td>
<td>No oscillation</td>
</tr>
<tr>
<td>3) Changes in host breadth spur</td>
<td>Papilionoidea (2573 spp.)</td>
<td>Hardy &amp; Otto 2014</td>
<td>Use of BiSSEness to test whether change in host breadth is associated with</td>
<td>Ø</td>
<td>Musical chairs</td>
</tr>
<tr>
<td>diversification or speciation</td>
<td></td>
<td></td>
<td>speciation events (speciational model vs gradual model of evolution for host</td>
<td></td>
<td></td>
</tr>
<tr>
<td>events are associated with shift</td>
<td></td>
<td></td>
<td>breadth)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>away from polyphagy</td>
<td>Nymphalini (31 spp.)</td>
<td>Janz et al. 2001</td>
<td>Estimate whether polyphagy is ancestral, estimate gains of new host vs</td>
<td>√</td>
<td>Oscillation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>losses of hosts (if polyphagy drives speciation; gains should exceed losses)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nymphalini (172 spp.)</td>
<td>Hardy 2017</td>
<td>Use of DEC model to reconstruct the history of host use and estimate gains</td>
<td>Ø</td>
<td>Speciation not associated with shift away from</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and losses of plants</td>
<td></td>
<td>polyphagy, no oscillation</td>
</tr>
<tr>
<td>4) Detoxification mechanisms in</td>
<td>Blepharidae (23 spp.)</td>
<td>Becerra 1997</td>
<td>Comparison of insect phylogeny with chemical defense similarity dendrogram</td>
<td>√</td>
<td>Coevolutionary arm race</td>
</tr>
<tr>
<td>insects evolve in response to plant</td>
<td></td>
<td></td>
<td>of host plants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>association</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blepharidae (37 spp.)</td>
<td>Becerra 2003</td>
<td>Comparison of timing of acquisition of plant defenses and insect</td>
<td>√</td>
<td>Coevolutionary arm race</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>counter-defenses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td>Author(s)</td>
<td>Year</td>
<td>Study</td>
<td>Verdict</td>
<td>Notes and Details</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Blepharidae (37 spp.)</td>
<td>Becerra</td>
<td>2003</td>
<td>Comparison of timing of acquisition of plant defenses and insect counter-defenses</td>
<td>✔</td>
<td>Coevolutionary arm race</td>
</tr>
<tr>
<td>Lygaeinae (20 spp.)</td>
<td>Bramer et al.</td>
<td>2015</td>
<td>Reconstruction of the ability to resist or sequester cardenolides</td>
<td>✔</td>
<td>Adaptation of insects to host plant defenses</td>
</tr>
<tr>
<td><em>Heliconius</em></td>
<td>de Castro et al.</td>
<td>2018</td>
<td>Review on correlated evolution of anti-herbivory adaptations in plants and counter-adaptations in <em>Heliconius</em></td>
<td>✔</td>
<td>Adaptation of insects to host plant defenses</td>
</tr>
<tr>
<td>Melitaenii (77 spp.)</td>
<td>Wahlberg</td>
<td>2001</td>
<td>Reconstruction of host association on insects and the presence of glycosides in associated plants</td>
<td>✔</td>
<td>Insects switch to chemically similar plants</td>
</tr>
<tr>
<td>174 taxa</td>
<td>Endara et al.</td>
<td>2017</td>
<td>Comparison of Lepidoptera assemblages associated with different plant species</td>
<td>✔</td>
<td>Similarity of assemblages on chemically similar plants; host association driven by similarity of plants</td>
</tr>
<tr>
<td>Apocynaceae</td>
<td>Livshultz et al.</td>
<td>2018</td>
<td>Evolution of cardenolide production in plants</td>
<td>✔</td>
<td>Plant defenses evolve in response to herbivorous insect predation</td>
</tr>
</tbody>
</table>

5) Diversification rates in phytophagous insect clades > non phytophagous insects

<table>
<thead>
<tr>
<th>Family</th>
<th>Study</th>
<th>Organisms</th>
<th>Study Description</th>
<th>Verdict</th>
<th>Notes and Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 families in various orders</td>
<td>Mitter et al.</td>
<td>1988</td>
<td>Sister groups comparison of phytophagous vs non phytophagous clades</td>
<td>✔</td>
<td>Phytophagy promotes diversification</td>
</tr>
<tr>
<td>1900 spp</td>
<td>Hunt et al.</td>
<td>2007</td>
<td>Sister groups comparison of phytophagous vs non phytophagous clades and estimation of diversification rates</td>
<td>Ø</td>
<td>Various types of niche shifts explain beetles diversification</td>
</tr>
<tr>
<td>Erolytidae, 53 taxa</td>
<td>Leschen &amp; Buckley</td>
<td>2007</td>
<td>Correlated evolution between species richness and phytophagy</td>
<td>Ø</td>
<td>Some shifts associated with phytophagy others not</td>
</tr>
<tr>
<td>367 spp, 172 families</td>
<td>McKenna et al.</td>
<td>2015</td>
<td>Infer shifts in diversification rates (MEDUSA)</td>
<td>✔ &amp; Ø</td>
<td>Different ecological factors prevail at different scales: phytophagy promotes diversification overall but not in all orders</td>
</tr>
<tr>
<td>31 insect orders</td>
<td>Wiens et al.</td>
<td>2015</td>
<td>Phylogenetic regression</td>
<td>✔ &amp; Ø</td>
<td></td>
</tr>
</tbody>
</table>

6) Major host shifts spur diversification

<table>
<thead>
<tr>
<th>Family</th>
<th>Study</th>
<th>Organisms</th>
<th>Study Description</th>
<th>Verdict</th>
<th>Notes and Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nymphalidae (54 genera)</td>
<td>Nylin &amp; Wahlberg</td>
<td>2008</td>
<td>Estimation of diversification rates after two major host shifts</td>
<td>✔</td>
<td>Escape and radiate and/or Oscillation</td>
</tr>
<tr>
<td>Butterflies (15 groups)</td>
<td>Fordyce</td>
<td>2010</td>
<td>Test for shifts in diversification rates across the phylogeny (LASER)</td>
<td>✔</td>
<td>Burst of diversification concomitant to some host shift, Escape and radiate</td>
</tr>
<tr>
<td>Adelpha (38 spp.)</td>
<td>Mullen et al.</td>
<td>2011</td>
<td>Test for shifts in diversification rates across the phylogeny (SymmeTREE)</td>
<td>✔</td>
<td>One rate shift attributed to hot shifts but also other ecological factors</td>
</tr>
<tr>
<td>Adelpha &amp; Limenitis (200 spp.)</td>
<td>Ebel et al.</td>
<td>2015</td>
<td>Test for shifts in diversification rates across the phylogeny (BAMM)</td>
<td>✔</td>
<td>Shift to Rubiacea played a role in insect diversification</td>
</tr>
<tr>
<td>Nymphalidae (398 genera)</td>
<td>Pena &amp; Espeland</td>
<td>2015</td>
<td>Detection of shifts in diversification rates (MEDUSA &amp; BISSE with trait= feeding on a particular plant family)</td>
<td>✔ &amp; Ø</td>
<td>Shift to Solanaceae spurred diversification</td>
</tr>
<tr>
<td>Hesperidae (290 genera)</td>
<td>Sahoo et al.</td>
<td>2017</td>
<td>Detection of shifts in diversification rates (LASER), BAMM, BISSE (trait= feeding on monocotyledonss vs dicotyledons) &amp; HISSE</td>
<td>✔ &amp; Ø</td>
<td>Several diversification shifts however could be associated with grassland expansion and geographic factors</td>
</tr>
<tr>
<td>Number</td>
<td>Note</td>
<td>Taxon (Number of spp)</td>
<td>Author et al. Year</td>
<td>Details</td>
<td>Notes</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-----------------------</td>
<td>--------------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>7)</td>
<td></td>
<td><strong>Pierinae</strong> (60 spp.)</td>
<td>Wheat et al. 2007</td>
<td>Distribution of a defense mechanism in insect and assessment of the glucosinolate in their host / comparison of rates of diversification in two sister clades with and without this defense</td>
<td>Escape and Radiate</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Pieridae</strong> (96 spp.)/Brassicales</td>
<td>Edger et al. 2015</td>
<td>Distribution of a defense mechanism and estimation of shifts in diversification rates in plants and insects (MEDUSA)</td>
<td>Complex coevolutionary arm race involving gene duplication in plants and associated with shift in diversification</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Asclepias</strong></td>
<td>Agrawal et al. 2009</td>
<td>Comparison of various model of evolution for plant defences including a speciation model</td>
<td></td>
</tr>
<tr>
<td>8)</td>
<td></td>
<td><strong>Erebia</strong> (74 spp.)</td>
<td>Pena et al. 2015</td>
<td>Test for shifts in diversification rates (BAMM &amp; DDD)</td>
<td>Density dependant model fits the data, diversification shifts not always significant depending on methods: acceleration of diversification associated to colonization of new area</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Cinara</strong> (92 spp.)</td>
<td>Meseguer et al. 2015</td>
<td>Test for shifts in diversification rates (TreePar)</td>
<td>Constant rate of diversification, no adaptive radiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Blepharoneura</strong> (49 spp.)</td>
<td>Winkler et al. 2018</td>
<td>Test for shifts in diversification (LASER, DDD)</td>
<td>Constant rate of diversification, no adaptive radiation, diversification patterns mostly explained by geographical factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Sesamiina</strong> (241 spp.)</td>
<td>Kergoat et al. 2018</td>
<td>Test for shifts in diversification through different methods (BAMM, *SSE)</td>
<td>Inverse patterns of diversification in insects and associated plants; host plant diversity alone does not explain insect diversification</td>
</tr>
<tr>
<td>9)</td>
<td></td>
<td>115 spp.</td>
<td>Farrell &amp; Mitter 1998</td>
<td>Sister-clade comparison of angiosperm vs non angiosperm feeding groups</td>
<td>Host-driven speciation</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Nymphalidae</strong> (309 genera)</td>
<td>Janz et al. 2006</td>
<td>Sister clade comparisons</td>
<td>Oscillation hypothesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Lymantrinae</strong> (55 genera)</td>
<td>Wang et al. 2017</td>
<td>Phylogenetic regression</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Cocidae</strong> (158 genera)</td>
<td>Lin et al. 2015</td>
<td>Regression without phylogenetic correction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Cecidomyiid</strong> (352 genera)</td>
<td>Joy &amp; Crespi 2012</td>
<td></td>
<td>Host-driven speciation</td>
</tr>
<tr>
<td>10) Clades including generalists speciate faster than clades with only specialists; diversification rates are positively correlated with host breadth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sister clade comparison of species richness between clades with only specialist and clades that encompass species with large host breadth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broadening of host range: plasticity facilitates the capture of new hosts and subsequent ecological speciation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clades experiencing higher rates of diversification experience a polyphagous state: oscillation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musical chairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11) Diversification rates of clades with labile host association&gt; diversification rates of clades with conservatism in host association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of *SSE to test whether speciation rates vary between lineages, correlation between host switching rates and diversification rates</td>
</tr>
<tr>
<td>Musical chairs</td>
</tr>
</tbody>
</table>

DEC (Ree, Smith 2008); Laser (Rabosky 2006); Medusa (Alfaro et al. 2009); BiSSE (Maddison, Midford, Otto 2007); BAMM (Rabosky et al. 2014); DDD (Etienne et al. 2012); HiSSE (Beaulieu, O’Meara 2016); SymmeTREE (Chan, Moore 2002)
Linnen et al. (2010) adopted yet a different approach to investigate the role of host shifts in speciation. They suggested that if host-shifts triggered speciation events, the evolution of host association should follow a speciational model of evolution, in which changes in host use occur during speciation events and its probability is not related to branch length. They thus compared the likelihood of a speciational vs a gradual model of evolution on a phylogeny of Neodiprion (Hymenoptera) and demonstrated that the speciational model was more likely, implying that host shifts accompany speciation events.

Hence, the studies that estimated niche differentiation at speciation events gave mixed support for scenarios where specialization onto different sets of host-plants is the main speciation process (Table 2.1). We must also keep in mind that these methods probably overlook many host shifts (shifts that resulted in population extinction do not leave any trace on phylogenetic trees) and therefore, these studies probably represent an overestimation of the impact of host shifts in speciation. It is nevertheless surprising that such studies have not been conducted on more study systems. This is perhaps likely due to the fact that they require a precise knowledge of the range of host-plants used by each insect species.

II.2.2) Host-breadth evolution

In order to test the predictions of macroevolutionary scenarios and investigate the role of specialization in insect species formation, phylogenetic studies have also investigated how host breadth is distributed throughout the evolutionary history of insects groups.

According to the “Oscillation hypothesis”, host breadth should vary a lot along the phylogeny of insects and the character state “generalist” should be transient (Nylin et al. 2014); i.e. the phylogenetic reconstruction of this character should show many transitions between specialists and generalists. Kelley & Farrell (1998), using ancestral character state reconstructions, have shown that host-breaths were indeed labile in Dendroctonus beetles and that specialists could evolve towards generalists. Several studies have then estimated the rates of transition between specialists and generalists. Using the phylogenies of fifteen insect genera from various orders, Nosil & Mooers (2005) estimated that the transition rates toward specialization exceed the transition rates toward generalization and that specialization was not a dead-end. More recently, Simonsen et al. (2010) reported repeated broadenings of diet in a genus of butterflies but did not quantify the phylogenetic signal of host breadth. Janz et al. (2001) and Nylin et al. (2014) have Many studies have indeed demonstrated the lability of host breadths (Table 2.2), which in Nymphalids using indicators of phylogenetic signal and
Wang et al. (2017) also showed that this character was highly labile in moths. This is compatible with the Oscillation hypothesis but does not necessarily prove it.

These patterns suggest that changes in host breadths are recurrent in the evolutionary history of insect lineages and are therefore compatible with the Oscillation hypothesis. In order to test whether these changes are linked with the colonization of new hosts (favouring subsequent specialization and speciation, some studies have investigated how gains and losses of new hosts along the phylogenetic history of lineage relate) to the number of speciation events - obtaining alternative results on Nymphalidae (Hardy, 2017 #4976). Janz et al. (2001) inferred the number of gains and losses of host plants throughout the history of Nymphalini. They found that gains exceeded losses and suggested that these were the result of repeated range expansions. However Hardy (2017) recently reanalysed the same dataset using different models of evolution of host use and inferred equal numbers of gains and losses of plants throughout the phylogeny of Nymphalini. He concluded that in this insect group host shifts and speciation events do not necessarily result from host range expansions and contractions. May-be another way to investigate the link between ‘host-breadth changes’ and ‘capture-gains of new hosts’ without using diversification analyses would be to investigate test for the correlated evolution of these two characters; i.e. explore how often the gains of new hosts is associated with a transition from specialist to generalists. In any case, the recent debates around the oscillation hypothesis (its predictions and how to test them) clearly demonstrates that ancestral character state reconstructions and statistical tests based on these reconstructions are very sensitive to model choices and character coding strategies and should always be interpreted with care, even when very specific predictions are verified.

II.2.3) Evolution of host-plant adapted traits

Many studies that aimed at finding support for the coevolutionary arm race hypothesized by the Escape and Radiate scenario have investigated the evolution of detoxification mechanisms in insect lineages and how they showed that they correlate with changes in their host associations of defences (Table 2.4). Studies that show that host switches occurred between plants with similar defenses, rather than closely related plants generally fit their results with the Escape and radiate hypothesis and Becerra (1997; 2003) reconstructed the evolution of detoxification mechanisms and plant defence in a genus of Chrysomelidae and their associated host plants (Burseraceae), and showed synchronous evolution of defences and counter defences that agree with a scenario of host-plant driven radiation in these insects. Similarly, in a review on the interaction between the chemically-
defended *Heliconius* butterflies and their *Passiflora* host-plants, de Castro et al. (2018) highlighted a large variety of anti-herbivory adaptations in the plants (including chemical defences, trichomes, fake eggs or herbivore damages, interaction with ants) and counter-adaptations in the butterflies that support a long history of coevolution. They conclude to coevolutionary arm race between insects and associated plants.

Also in agreement with the *Escape and Radiate* theory, Wahlberg (2001) using ancestral character states reconstruction on a phylogeny of Melitaeini (Lepidoptera) showed that host switches occurred between plants with similar defences, rather than closely related plants (i.e. host-plant use was more influenced by chemistry than by plant taxonomy). Endara et al. (2017) showed opportunistic host switches to plants with similar defences in a genus of butterfly. They however some of these studies argue in favour of a scenario in which suggested that herbivores simply "choose" host-plants based on their own defensive traits. The authors conclude that disruptive selection is not a driver of speciation in this case. Many other studies looked at the diversification dynamics of genes involved in plant detoxification and how it correlates with changes in host-plant associations. Wheat et al. (2007) showed that the evolution of a detoxification mechanism, a nitrile-specifier protein (NSP) in Pieridae matches the distribution of glucosinolate in their host-plants. Edger et al. (2015) further investigated arm races between Brassicales and butterflies (Pieridae) and showed that repeated evolution of nitrile-specifier proteins were associated with bursts of diversification over the past 80 Myr in Pieridae. Calla et al. (2017) examined the cytochrome P450 monooxygenase (CYP) gene superfamily diversification in the genomes of seven Lepidoptera species varying in host-breadth. They showed that its dynamics (duplication and losses) was correlated with the ability to metabolise defences. Bramer et al. (2015) analysed the ability to sequester toxic cardenolides throughout the phylogeny of the hemipteran subfamily Lygaeinae and suggested that it was acquired in response to selection by cardenolide-producing Apocynaceae host-plants. Host switches underlie insect species differentiation but this differentiation is not the product of coevolutionary interactions.

Reconstruction of the history of plant defences were also conducted on plant phylogenies. For instance, Livshultz et al. (2018) reconstructed the evolution of cardenolide production in Apocynaceae and suggested that it could have evolved in response to herbivorous insect predation. On the other hand, Agrawal et al. (2009), reconstructed the history of plant defences in North American milkweed species (*Asclepias*, Apocynaceae) and showed that less investment in cardenolide production correlates with an increase in speciation rates. This does not follow the predictions of “*Escape and Radiate*”...
direct tests of acceleration of diversification upon the acquisition of new defense mechanisms rather sometimes contradict the ‘Escape and Radiate” prediction and suggests that investment in costly defenses might have impeded plant diversification in this plant group.

In contrast to detoxification mechanisms, traits (and underlying genes) involved in host recognition and host-plant choice (chemosensory traits) have been less studied in a phylogenetic context. However their evolutionary dynamic probably plays as important a role as adaptations to plant defenses in phytophagous insect speciation (Smadja, Butlin 2009). (Matsuo 2008) showed that an odor binding protein in 27 Drosophila species can evolve relatively fast in closely related species through gene duplications and losses, and proposed that this dynamic could explain the evolution of host preferences in this species complex.

Sánchez-Gracia et al. (2009) and Vieira & Rozas (2011) conducted a comparative genomic analysis of odor binding protein and chemosensory proteins from the genomes of several Arthropoda species (mainly Drosophila). They showed a high number of gains and losses of genes, pseudogenes, and independent origins of gene subfamilies. This dynamic, if analyzed in relation to host choices and host breadth in a phylogenetic context, could explain some host shifts and subsequent speciation events. Focusing on behavioral traits of insects, Molnar et al. (2018) analyzed the antennal responses of 12 gall midge species to a wide range of host-plant-related volatiles and showed that species with similar response shared host-plants. Finally, deciphering the evolutionary dynamics of genes involved in mate recognition and their link with host association could also inform us on the role of host-plants in the speciation of insects that feed (and often mate) on them. For instance, (Schwander et al. 2013) showed that cuticular hydrocarbon profiles involved in mate choices vary among Timema species (Coleoptera Phasmatodea), and that most evolutionary changes in hydrocarbon profiles occur in association with host-plant shifts and speciation events in this genus of phytophagous insects. This study shows that physiological traits involved in reproductive isolation can be associated with host plant differentiation on a macroevolutionary time scale.

In summary, many studies have investigated the evolution of detoxification mechanisms in insects butterflies and found support for framing their hypotheses within the Escape and Radiate theory. More studies investigating changes in traits involved in host recognition throughout the diversification of insects are rare and they are generally not framed within host-plant driven scenarios. Such studies are needed in order to determine which traits underlie host-plant associations and whether their evolution accompany drives speciation events (see Perspectives).
II. 3) Studying how host-plants use impact the diversification dynamics of herbivorous insects

Methods for testing diversification dynamics have expanded over the last decade. Testing for the existence of temporal bursts of diversification was once restricted to analyses of groups with comprehensive fossil records. Diversification dynamics can now be studied through comprehensive phylogenies (Goswami, Mannion, Benton 2016). Given the breadth of available methods, theoretically, nearly all hypotheses can be put to test.

II.3.1) Are phytophagous insects more diversified than their related counterparts?

Studies that posit that host-plant adaptation favours phytophagous insect diversification predict that those are much more diversified than non-phytophagous insects. However this assertion deserves to be statistically tested. The first studies addressing this question (Farrell & Mitter 1990; Farrell & Mitter 1998; Mitter et al. 1988), compared the diversity (number of species) of phytophagous vs related non-phytophagous clades in beetles. They all suggested that herbivorous clades are more diverse than their non-phytophagous sister clades (Table 2.5). However, Hunt et al. (2007) and then Rainford et al. (2015), using more comprehensive beetle phylogenies associated with up-to-date recent comparative methods, and also adopting a sister-clade comparison, did not find a significant increase in species richness in phytophagous clades. In their phylogeny of the beetle family Erotylidae, Leschen & Buckley (2007) mapped the evolution of species number within clades (as a two state character: 0/1) and did not detect any significant correlation between species richness and phytophagy (using Bayesian tests of character correlation). McKenna et al. (2015) tested for temporal variation in diversification rates (using MEDUSA Alfaro et al. 2009) on a global phylogeny of Coleoptera and showed that those underwent several accelerations of diversification rates; some seemed to be associated with the evolution of phytophagy while others were not suggest that several types of niche shifts beside phytophagy can explain insects species richness.

There are more informative tests than those that merely test whether one group contains more species or has diversified more rapidly than others. More recently Wiens et al. (2015) used Phylogenetic Generalized Least Square Regressions and comprehensive phylogenies within and between several insect orders and showed that the proportion of phytophagous insect species in a clade was positively correlated with diversification rates in several insect orders. But this relationship did not hold for Coleoptera, Hymenoptera and Orthoptera.

Studies within
Insect orders have mostly been addressed within Coleoptera – as they are by far the most diverse order but more in-depth analyses of other orders would actually also be useful. To our knowledge, likelihood-based character state dependent diversification models (known as the \( \text{--}_S \text{SSE} \) models, such as BiSSE, ClaSSE and QuaSSE for binary, multistate and quantitative traits, respectively (Maddison, Midford, Otto 2007; FitzJohn 2010; Maddison, FitzJohn 2015), which can specifically test whether transition in character states are associated with variations in speciation and extinction rates. These diversification models can be used to test whether the evolution of phytophagy has favoured diversification. However, they were not yet used in Wiens et al. (2015) to test the role of phytophagy in insect diversification. A robust estimate of diversification parameters by these methods would actually require dense and random species sampling throughout the phylogeny. According to (FitzJohn 2010) more than 15% of the species must be included in the phylogeny in order to conduct unbiased \( \text{--}_S \text{SSE} \) tests. The study was conducted at the order level and within each order the proportion of missing species was too high to apply \( \text{--}_S \text{SSE} \) methods.

From current studies using more comprehensive phylogenies and statistical tests, it is thus yet not completely clear whether phytophagy increases diversification rates in insects in comparison to other life habits. Therefore character state dependent diversification models await The availability of more comprehensive phylogenies will allow testing this hypothesis on-in more insect orders using character state dependent diversification models. Nevertheless; a caveat of these diversification models is that they may overlook more complex models involving many unmeasured and co-distributed traits (especially when few transitions in feeding habits are observed). In particular, for phytophagy, other traits that may drive shifts in diversification may cause a spurious detection of increased diversification rates in phytophagous insects if this trait’s states partly correlate with phytophagy states, or may instead erase any signal due to phytophagy. The HiSSE model (Beaulieu, O’Meara 2016) which models hidden characters that might influence diversification might help untangling these confounding factors.

II.3.2) Testing for adaptive radiation

Diversification analyses can also be applied to test whether the capture of new host-plants has favoured adaptive radiations. Under such a scenario, the diversification curve of phytophagous insect clades should exhibit early bursts of speciation upon the capture of new
groups of host-plants. Eventually, insect lineages are eventually expected to fill the newly found niche space and the diversification curves should reach a plateau.

There are several studies that have observed acceleration in diversification—speciation rates in insect lineages and have put forward ad hoc narratives that attribute this acceleration to the capture of new host-plant lineages (Table 2.6) or detoxification mechanism (Table 2.7). These studies have mainly been conducted ion butterflies. While most studies have observed shifts in speciation rates in insects upon host plant acquisition, for instance, Wheat et al. (2007) showed that the Pierinae clade that can detoxify glucosinolate in their host plants are more diversified that its sister clade. Mullen et al. (2011) have observed an increase in species richness in the butterfly genus Adelpha (Nymphalidae) and attributed this increase to adaptive divergence in response to host plant diversity found within the neotropics (namely host-plant shift to Rubiaceae and other plant families). Ebel et al. (2015) revealed an increase in the rates of diversification on the phylogeny of the genus Adelpha and observed that those were concomitant to host shifts. Sahoo et al. (2017) revealed two accelerations in diversification in skipper butterflies that they attributed to shifts from dicots to monocots at time where those diversified and expanded—their results seem to be sensitive to the models used. Studies have also rarely included trait-dependent diversification models and when they did, they often relied on single or few colonization events that were also correlated with important geographical changes, making it difficult to conclude on the causal effect of the sole host shift (Table 2.6). Fordyce (2010) showed changes in diversification rates upon the capture of new plant lineages in several butterfly lineages and Edger et al. (2015) detected shifts in diversification rate associated with the colonization of new host-plants with new defences, though this was not statistically tested. Winkler et al. (2018) on the other hand, uncovered stable diversification rates and no decrease associated with overlap in host use throughout the phylogeny of the fly genus Blepharoneura, suggesting that this lineage did not undergo an adaptive radiation. Using BiSSE, Peña & Espeland (2015) found that a hostplant shift to Solanaceae was correlated with an increase in net diversification rates in Ithomiini butterflies (Nymphalidae). However, since only one shift to Solanaceae occurred in the Nymphalidae, this correlation should not be taken as evidence that hostplant shift has driven diversification of this butterfly tribe. Instead, the radiation of Ithomiini may be linked with geographical context in this study.

While several studies have uncovered acceleration in diversification rates upon the capture of new host-plant lineages. Among these studies addressing investigating adaptive
radiation patterns (Table 2.8), only a handful have investigated whether the number of species reached a plateau after an initial burst of diversification. Meseguer et al. (2015) studied the diversification dynamics of a conifer-feeding aphid genus. They revealed an accelerated rate upon the capture of conifers but no saturation as expected under a scenario of adaptive radiation. This is better addressed using The use of density-dependent models of diversification (DDD) would be better suited to test whether rates of diversification decrease through time during evolutionary radiations once the number of species supposedly reach the carrying capacity of the novel ecological niche (Rabosky, Lovette 2008; Etienne et al. 2012). But these tests can only be applied to lineages in which the number of species in each subclade is well known. Peña et al. (2015) used such diversity-dependent birth-death models to infer speciation, extinction rates and carrying capacity on a phylogenetic tree of the butterfly genus Erebia and showed that the diversification dynamics was consistent with a model of adaptive radiation. In any case, studies generally provided weak support for adaptive radiation scenarios (Table 2.8). They all suggested that Kergoat et al. (2018) compared the diversification dynamics of Sesamiina stemborer moths and their associated grasses. The initial burst of diversification observed in moths suggests that their emergence might have been favoured by the diversification of their host-plants. A decline in speciation rates was then inferred in these insects using Diversity-dependent birth-death models. However, their host-plants continued to thrive. This challenges the “adaptive radiation scenario” and suggests that these moth diversification patterns cannot solely be explained by the availability of suitable host-plants and insisted that abiotic factors such as geography and temperature should be taken into account in diversification scenarios. (Kergoat et al. 2018).

II.3.3) Correlating host breadth with diversification dynamic

Advocates of the Oscillation hypothesis suggest that clades showing a higher diversity of host-use (using more host-plant species altogether) should be more diverse than their sister clade (Janz, Nylin, Wahlberg 2006), and this prediction has been verified in butterflies, Coccidae and some gall inducing flies (Table 2.9). Janz et al. (2006) and Nylin & Wahlberg (2008) found a positive relationship between species diversity of butterflies and species diversity of host-plant taxa in several genera of Nymphalidae, through a sister clade comparisons of number of species. Joy & Crespi (2012) and Lin et al. (2015) found a similar result in respectively gall-inducing flies and Coccidae. Wang et al. (2017) used a sister clade comparisons and PGLS to demonstrate a similar trend in moths.
Although these results support a model where the diversity of phytophagous insects is sustained by the diversity of the hosts they use, they could fit both a model of Oscillation and the Musical chairs hypothesis. The latter indeed predicts that lineages including specialist species that often switch between hosts use a large number of host-plants. To tell apart the Oscillation from the Musical chairs, it is actually more informative to test how host breadth variations affect insect diversification dynamics. According to the Oscillation hypothesis, clades including generalist species should be more speciose than clades including only specialists—In order to test this prediction, diversification analyses have been conducted on various insect phylogenies. Early papers have used fairly basic methods such as ‘sister clade analyses’ while Weingartner et al. (2006) tested this prediction in Polygonia butterflies (Nymphalidae) through sister-clade analyses and showed that clades using a broader host-plant range are more species-rich than their sister group that encompass species that only use the ancestral hosts (here, urticalean rosids), in agreement with the Oscillation scenario.

More finely tuned analyses, such as more recent investigations have adopted trait-dependant diversification models (BiSSE and QuaSSE models) (Table 2.10) have been used on the phylogenies of Papilionidae and the tribe Heliconini by Hardy and Otto (2014). In this paper, where the Musical chairs hypothesis was put forward, the authors showed that: 1) speciation rates decreased with host breadth (i.e. monophagous species diversified faster than polyphagous species); 2) changes in host breadth were not associated with cladogenetic events; 3) rates of host switching were positively associated with diversification rates. All these results were in agreement with the Musical chair hypothesis. Still on Nymphalidae, Hamm & Fordyce (2015) found that host breadth characteristics were phylogenetically conserved which does not fit the predictions of the Oscillation hypothesis about the lability of host-breadth and also found that speciation rates did not increase with host-breadth expansion, in agreement with the Musical chair scenario.

By contrast, Hardy et al. (2016) showed on a Coccidea phylogeny that host breadth (measured as the number of host-plant families) was positively correlated to speciation rates. As SSE models are known to generate false positives (Bouchenak-Khelladi et al. 2015; Rabosky & Goldberg 2015; Davis et al. 2013), they conducted the analysis on a set of randomized trees in order to test whether the constrained diversification model (i.e. the model in which evolutionary transitions in character states are associated with shifts in extinction and/or speciation rates) is also chosen in these analyses. The authors did not frame their study within the Oscillation vs Musical chairs controversy, but if we refer to the paper of Hardy & Otto (2014), these results fit with some of the predictions of the Oscillation hypothesis.
These studies provided support for alternative scenarios. But even when a specific prediction was met, the authors could not always reach a firm conclusion. As already previously mentioned, there are debates on the influence of host breadth variation on diversification dynamics (Hamm & Fordyce 2015; Janz et al. 2016; Hamm & Fordyce 2016; Wang et al. 2017). The transient nature of the generalist feeding diet under host-driven speciation scenarios makes it difficult to derive clear predictions. The negative or positive relationship observed between host breadth dynamics and diversification can always be obscured by the rapid evolution of host breadth (Nylin, Janz 2009; Janz et al. 2016). The way host breadth is measured (binary vs continuous) is also known to affect the results (Hamm, 2015) and it has been long recognized that categorizing species into either specialist or generalist can be difficult and somewhat subjective (Janz, 2001). Furthermore, although the musical chair yields specific predictions, those are mainly rebuttals of the “Oscillation” predictions (but see Table 2.11), but rejecting an “Oscillation” scenario does not instantly mean that a “Musical chair” scenario is at play. Hence despite the existence of sophisticated methods and comprehensive phylogenies, the prevalence of these scenarios in insects is still debated.

Furthermore, another caveat of these studies lies in the distinction between speciation and extinction rates. Indeed, if specialization towards host plants can accelerate speciation rates it can also increase extinction risks when plants are not highly abundant. Although diversification methods can potentially differentiate extinction rates from speciation rates (Morlon 2014). Consequently, predicting exactly how changes in host-plant breadth affects diversification dynamics remains difficult. Finally, diversification methods such as *SSE models are known to generate false positives (Bouchenak-Khelladi et al. 2015; Rabosky, Goldberg 2015; Davis, Midford, Maddison 2013), and can thus inflate the role of host plant and host breadth in diversification when those are investigated. It is therefore, highly recommended to conduct the analyses on a set of randomized trees in order to test whether the constrained diversification model (i.e., the model in which evolutionary transitions in character states are associated with shifts in extinction and/or speciation rates) is also chosen in these analyses (as done in (Hardy, 2016)).

In summary, diversification analyses have been widely applied to phytophagous insects. These methods were fairly basic in early papers but they have been refined throughout the years and their use allowed reframing hypotheses on the role of host plant specialization on insect speciation. Recent results brought mixed evidence for phytophagy as an accelerator.
of diversification. Trait-dependant diversification analyses supported alternative scenarios involving oscillation in diet breadths as a driver of host-plants shifts and speciation, and from our review, it is likely that different scenarios will prevail in different lineages and even probably in the same lineage at different time scales.

Finally Moving away from methodological issues, our survey also underlines that many studies that explore macroevolutionary scenarios have been conducted on butterflies Lepidoptera. In order to have a better understanding of how phytophagous insects capture new host-plants and whether it influences the role of host-plant shifts in insects their diversification dynamics, it seems necessary to test the predictions of macroevolutionary scenarios on other insect groups. Aphids (Hemiptera) could be good candidates for such investigation. Their range of host-plants is very well documented (Blackman, Eastop 2006; Holman 2009). Although most aphids are host-specific, there are some polyphagous species. Some species are even only polyphagous during the asexual part of their life-cycle (Moran 1992; Jousselin, Genson, Coeur d'Acier 2010; Hardy, Peterson, von Dohlen 2015). This temporary broadening of diet has already been suggested to facilitate the capture of new host-plants (Moran 1992) and subsequent specialization and speciation of these new hosts. These life-cycle y-transitions could have thus favoured the diversification on this group clade (Moran 1992; Jousselin, Genson, Coeur d'Acier 2010). This scenario fits the Oscillation hypothesis and should be tested on a robust and comprehensive aphid phylogeny. Other insect Hemiptera groups for which host plant repertoire are is well known such as Coccidae (Garcia-Morales et al. 2016) and Psylidae psyllids (Ouvrard et al. 2015) could also be used to test the predictions of on the role of host breadth evolution in macroevolutionary scenarios. The limitation lies in the availability of robust and comprehensive phylogenies for these groups.

In summary, diversification analyses have been widely applied to phytophagous insects. These methods were fairly basic in early papers but they have been refined throughout the years and their use allowed reframing hypotheses on the role of host plant specialization on insect speciation. Recent results brought mixed evidence for phytopagy as an accelerator of diversification. Trait dependant diversification analyses supported alternative scenarios involving oscillation in diet breadths as a driver of host plants shifts and speciation, and from our review, it is likely that different scenarios will prevail in different lineages and even probably in the same lineage at different time scales.

III Perspectives
As seen throughout this review, phylogenetic comparative methods provide the template to test hypotheses on the role of host plant association in the speciation of phytophagous insects. While those methods have undoubtedly advanced the field significantly since “the Escape and Radiate” paper, readers must keep in mind that phylogenetic comparative methods often rely on mere correlations. Significant associations between character changes and the cladogenetic events might arise as a consequence of speciation itself when post-speciational character changes occur. Furthermore, comparisons of models of evolution such as those used in trait-dependent diversification analyses often rely on trees that encompass few transitions in character states and are therefore not always robust \( (\text{see } \text{Beaulieu, 2013 #5053}) \). In such analyses, the “best model” is not necessarily the true model and significant \( P \) values should not be interpreted as strong evidence for an evolutionary scenario. Finally, all these methods are very sensitive to sampling biases and those need to be carefully taken into consideration.

In addition to using the approaches focused on host plant associations and diet breadth cited throughout this review, one way to further investigate hypotheses of speciation driven by associations with host-plants would be to integrate a variety of data in a phylogenetic context. Below we outline three potential directions for future research: 1) disentangling the role of plant-insect interaction from that of co-variates, such as geography and climate; 2) combining phylogenetic analyses with interaction network approaches including other partners, at various ecological scales (from community-scale to global scale); and 3) studying traits and genes underlying the association; 3) combining phylogenetic analyses with interaction network approaches including other partners, at various ecological scales (from community-scale to global scale).

III.1) Investigating the role of abiotic factors: geography and climate

Geography and ecology are always closely intertwined in speciation scenarios. There have been several studies that have investigated geographic range expansion in herbivorous insects (Becerra, Venable 1999; Slove, Janz 2011); climate induced host shifts (see (Winkler, Mitter, Scheffer 2009; Nyman et al. 2012 for a review and recent studies since then {Lisa De-Silva, 2017 #4641; Sanchez-Guillen et al. 2016; Lisa De-Silva et al. 2017; Owen et al. 2017; Pitteloud et al. 2017; Sahoo et al. 2017}) and climate driven diversification dynamics (Kergoat et al. 2018). All these studies suggest that abiotic factors are entangled with host-plants changes in species diversification scenarios. However there are few studies that explicitly test the predictions of speciation through geographic isolation (Barraclough, Vogler 2000) and
whether these events systematically accompany host shifts or sustain most speciation events (but see Jordal & Hewitt 2004; Jousselin et al. 2013; Doorenweerd, van Nieukerken, Menken 2015; Hardy, Peterson, Normark 2016). Such analyses are important if we want to tell whether adaptations to new host plants represent post-speciational changes following geographic isolation rather than the main driver of speciation events. Cospeciation methods that take into account the biogeographic history of interacting lineages (Berry et al. 2018) could be also used to investigate whether host shifts are associated with dispersal events in systems where hosts plants and insects phylogenies show some congruent patterns.

## II]]

### 2. Combining phylogenetic with interaction network approaches, at various ecological scales

Herbivores and the plants they feed on form interaction networks, and as such the structures of the networks can be characterized by several parameters, such as modularity (the propensity of a group of species to interact with a similar set of partners) and nestedness (the propensity of specialist species to interact with generalist species and vice-versa). Antagonistic interaction networks, such as plant-herbivore networks, tend to be highly modular (Thébault & Fontaine 2010). A recent study combining interaction network with phylogenetic approaches on simulated and real datasets predicted that the *Escape and radiate* scenario should produce a modular network structure, whereas the *Oscillations* scenario should produce a more nested structure (Braga et al. 2018). When applied to real data (two butterfly families, Nymphalidae and Pieridae), this approach revealed that host-plant-butterfly networks tend to be both modular and nested, which the authors interpret as being the result of a complex pattern of diversification, involving both episodes of radiation on new hosts (producing modules containing closely related species) and occasional shifts to other host lineages, producing both nestedness within modules and connections between modules. Additionally, phylogenetic and network approaches could be expanded to encompass other interacting partners (e. g. Elías et al. 2013, Ives & Godfray 2006). Indeed, insect-host-plant communities can be seen as ecosystems where biotic interactions, such as parasitism and mutualism also take place (Forister et al. 2012). These other partners can indirectly influence the interaction between plants and there herbivores: e. g., direct competition (Jermy 1988) apparent competition between herbivores, stemming from shared natural enemies, (Holt 1977), and vice-versa (e. g., when herbivory elicits anti-herbivore defences mediated by
herbivore enemies, (Fatouros et al. 2008). Multitrophic interactions probably explain many
diversification patterns in herbivorous insects (Singer & Stireman 2005).

Finally, such approaches could be applied both at a large scale (e.g., Braga et al. 2018), to
embrace global patterns of diversification and interaction, or at the community level (Elias et al. 2013, Ives & Godfray 2006), where interactions actually occur, and where fine scale
processes (e.g., host-plant shift at the species or the population level) can be unveiled.

III.32) Unravelling traits involved in the interaction and their underlying genes, and
integrating this information in phylogenetic studies

Interactions between insects and their host-plant are ultimately mediated by traits, such as
host-plant defences and the capacity of circumventing plants defences, but also host-plant
cues and the capacity for herbivores to detect those cues. Characterizing such traits, their
genetic determinism and looking at their evolutionary trajectory would greatly advance our
understanding of the diversification of insects (e.g., (de Castro et al. 2018). Testing whether
different trait states are associated with different speciation rates can be performed using
*SSE methods (e.g., as in Onstein, 2017 for a trait relating to frugivory in palm trees). In
addition, methods that test whether patterns of trait evolution conform to a model accounting
for interactions mediated by those traits are currently being developed (Manceau, Lambert,
Morlon 2016) (Drury et al. 2017), and they could inform on the processes underlying
herbivore diversification. However, targeting traits involved in plant-insect interactions may
be challenging. Pivotal traits are difficult to identify, they include chemical, behavioral and
metabolic traits and when they are properly characterized they are often multigenic.

Perhaps a promising direction for future research is the implementation of a hybrid
 genomic approach that combines transcriptomics, phylogenomics, comparative analyses and
population genomics (see (Nevado et al. 2016). In such approaches, full transcriptomes of
species from a target clade (for instance, a clade of phytophagous insect) are generated. These
transcriptomes (or other sequence data) are used to generate a phylogeny, where classical
diversification and character evolution tests can be performed (evolution of characters, test for
diversity-dependent diversification, shifts in diversification following host-plant shift). Then,
genomes under selection can be detected from transcriptomic data using population genetics
statistics, and can be matched to existing databases (e.g., Lepbase for Lepidoptera, Challis et
al. 2016) for identification purpose. Additionally, genes that are down or upregulated can also
been detected by classical tests of differential expression and identified, and the association of
genes under selection, either via different sequence or expression pattern, with species
diversity can be tested. **The main limits of this approach lie** in the availability of specimens (transcriptomic data need to be obtained from fresh or suitably preserved tissues; biological replicates are needed), those might be difficult to obtain throughout an entire phylogenetic tree and the availability and quality of the reference gene database to match genes with putative functions might also limit the applications of this approach.

### III.3) Combining phylogenetic with interaction network approaches, at various ecological scales

Herbivores and the plants they feed on form interaction networks, and as such the structures of the networks can be characterized by several parameters, such as modularity (the propensity of a group of species to interact with a similar set of partners) and nestedness (the propensity of specialist species to interact with generalist species and vice-versa). Antagonistic interaction networks, such as plant-herbivore networks, tend to be highly modular (Thébault, Fontaine 2010). A recent study combining interaction network with phylogenetic approaches on simulated and real datasets predicted that the *Escape and radiate* scenario should produce a modular network structure, whereas the *Oscillations* scenario should produce a more nested structure (Braga et al. 2018). When applied to real data (two butterfly families, Nymphalidae and Pieridae), this approach revealed that host-plant butterfly networks tend to be both modular and nested, which the authors interpret as being the result of a complex pattern of diversification, involving both episodes of radiation on new hosts (producing modules containing closely related species) and occasional shifts to other host lineages, producing both nestedness within modules and connections between modules. Additionally, phylogenetic and network approaches could be expanded to encompass other interacting partners (e. g. (Elias, Fontaine, van Veen 2013), (Ives, Godfray 2006)). Indeed, insect-host-plant communities can be seen as ecosystems where biotic interactions, such as parasitism and mutualism also take place (Forister et al. 2012). These other partners can indirectly influence the interaction between plants and their herbivores: e. g., direct competition (Jermy 1988) apparent competition between herbivores, stemming from shared natural enemies, (Holt 1977), and vice-versa (e. g., when herbivory elicits anti-herbivore defences mediated by herbivore enemies, (Fatouros et al. 2008)). Multitrophic interactions probably explain many diversification patterns in herbivorous insects (Singer, Stireman 2005). Finally, such approaches could be applied both at a large scale (e. g., Braga et al. 2018), to embrace global patterns of diversification and interaction, or at the community level (Elias, Fontaine, van Veen 2013), (Ives, Godfray 2006), where interactions actually occur, and where
fine-scale processes (e. g., host-plant shift at the species or the population level) can be unveiled.

Conclusions

Many phylogenetic studies of plant-insect associations now include formal tests of macroevolutionary scenarios involving host-driven speciation. In an attempt to summarize the literature on this topic, we show that the predictions of host-plant driven speciation are not straightforward and can vary depending on studies. We advocate a standardization of these predictions to facilitate cross study analyses. Furthermore, it is also recognized that different scenarios can leave the same phylogenetic signature (Janz et al. 2016) and that depending on the analytical approaches undertaken to test the predictions laid out in Table 1, conclusions can vary (Hardy et al. 2017). Unfortunately this means that the interpretations of phylogenetic inferences can remain somewhat subjective. But these shortcomings should not obscure the progresses that have been made in the field. Phylogenetic comparative analyses help framing hypotheses and clarify some of the narratives used to explain the diversification of phytophagous insects. In order to move towards a standardization of phylogenetic approaches, we propose here a (non-exhaustive) list of relatively simple tests that could be applied to an insect phylogeny that includes robust data on host plant association (Fig. 2). The limitations of these approaches (sampling biases, false positives) have been described throughout this review. Finally, this survey of the literature shows that: 1) the simple assumption that phytophagy has accelerated insect diversification is not always sustained by meta-analyses; 2) the expectation that sister lineage will use different ranges of host plants is not often tested, and, when it is, the predictions of a host-driven speciation scenarios are not always met. We then underline that the results of phylogenetic comparative methods cannot be interpreted as hard evidence as they remain mere correlations. In the end, a full understanding of the processes explaining the diversification of phytophagous insects will require the integration of phylogenies with other data sources and analytical methods. We propose here a few perspectives to integrate such data and investigate host-driven speciation scenarios on a macroevolutionary time scale.

Finally, to conclude, if the last two decades have seen the rise of molecular phylogenies and the development of analytical methods that include ecological data, this should not obscure the fact that these data needs to be thoroughly curated before any
phylogenetic comparative analyses. Qualifying host associations of insect species necessitates field work and advanced taxonomy, as mistakes can seriously impact the results of macroevolutionary studies. Functional studies aimed at deciphering host-plant adapted traits in insects (and in particular traits implied in host choice) and characterizing genes that underlie these traits are also needed to integrate this data in a phylogenetic context and link microevolutionary processes with macroevolutionary scenarios.
Figure 2:

a) Host-plant range history

- Range 1
- Range 2
- Range 3
- Range 4

H1: Speciation driven by host shifts
H2: Speciation not driven by host shift

Tests: evaluate the number of host shifts and compare models of evolution for host range.

Under H1: number of host shifts ≈ number of speciation events > number expected when host use is randomized; host-plant range follows a speciational model of evolution.

Under H2: host plant range exhibits a phylogenetic signal.

b) Host breadth history

- specialist
- generalist

H1: No oscillation
H2: Oscillation

Tests: evaluate phylogenetic signal and compare models of evolution for host breadth.

Under H1: host breadth exhibits a phylogenetic signal.

Under H2: host breadth follows a speciational model of evolution.

c) Diversification analyses

- specialist
- generalist

H1: No adaptive radiation
H2: Adaptive radiation driven by host shifts (Escape and radiate)
H3: Adaptive radiation driven by shifts in diet breadth (Oscillation)

Tests: use sister-clade analyses, PGLs, diversity dependant models of diversification or *SSE models to test the effect of host shift, host breadth and host plant lineage diversity on speciation.

Under H1: no significant shifts in speciation rates upon host shifts.

Under H2: significant shifts in speciation rates upon host shifts, number of insect species higher in species rich plant clades.

Under H3: speciation rate increase following shift away from polyphagy.
Suggestions of phylogenetic comparative methods that can be deployed to test the prediction of host-driven speciation scenarios. a) Tests that rely on a robust reconstruction of host-plant range evolution. b) Tests that rely a robust reconstruction of host-breadth. c) Diversification analyses that rely on reconstruction of both these characters. H1, H2, H3 are alternative scenarios and are represented by a schematic phylogenetic reconstruction above.
Glossary:

Adaptive radiation: the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage (Schluter 2000). It occurs when natural selection drives divergence of an ancestral species into descendants that exploit different ecological niches.

Coevolution: reciprocal evolutionary changes occurring in two or more species that result from reciprocal selective pressures exerted by the interacting partners.

Coevolutionary diversification: when diversification patterns arise from coevolution.

Cospeciation: simultaneous speciation events in lineages involved in long-term interspecific associations which result in congruent phylogenies and temporal congruence of speciation events.

Diversification dynamic: rates of species formation and extinction through time.

Ecological specialization: when species are limited themselves to a restricted set of resources (diet-habitat-niches), as a result of evolutionary trade-offs.

Evolutionary lability: tendency for a character to change frequently throughout evolution.

Generalists: species that use a wide niche (a wide range of host plants belonging to different lineages in the particular case of phytophagous insects).

Host-plant adaptation: heritable trait that confers a selective advantage on a particular host-plant.

Phylogenetic conservatism (-i; tendency for closely related species to be more similar than expected under Brownian motion evolution.

Specialists: species that use a narrow niche.

Phylogenetic tracking occurs when a host-dependent species (here phytophagous insects) diversifies to utilize niches created by the speciation its hosts (here host-plants), related
insects use related plants, this leads to parallel phylogenetic trees but no temporal congruence of speciation events

**Phytophagous insect:** an insect that feeds on any plant organ during whole or part of its life cycle, it excludes pollinators feeding on nectar and pollen but include pollinators that feed on developing seeds (*i.e.* seminiphagous insects).

**Specialists:** species that use a narrow niche (a restricted set of resources).

**Sister species/sister clades:** the closest relatives of another given unit (species/clade) in a phylogenetic tree.

**Conflict of interest disclosure**

The authors of this preprint declare that they have no financial conflict of interest with the content of this article.

**References**


Kergoat, GJ, BP Le Ru, SE Sadeghi, M Tuda, CAM Reid, Z Gyorgy, G Genson, CS Ribeiro-Costa, A Delobel. 2015. Evolution of Spermophasus seed beetles (Coleoptera, Bruchinae, Amblycineri)
indicates both synchronous and delayed colonizations of host plants. Molecular Phylogenetics and Evolution 89:91-103.


Vea, IM, DA Grimaldi. 2016. Putting scales into evolutionary time: the divergence of major scale insect lineages (Hemiptera) predates the radiation of modern angiosperm hosts. Scientific Reports 6:23487.


