

Evolution of host repertoires and the diversification of butterflies

Mariana Pires Braga

Academic dissertation for the Degree of Doctor of Philosophy in Animal Ecology at Stockholm University to be publicly defended on Friday 15 February 2019 at 10.00 in Vivi Täckholmsalen (Q-salen), NPQ-huset, Svante Arrhenius väg 20.

Abstract

All herbivorous insects are specialized to some extent to their host plants, but the level of specialization varies greatly. Insect-plant coevolution is often invoked to explain the large diversity of herbivorous insects, but the role of specialization during diversification is still controversial. Although well-studied, our understanding of the evolution of species interactions is still improving, and recent theoretical developments have highlighted the role of generalization (via colonization of new hosts) on diversification. In this thesis, various approaches are combined for a detailed study of the origins of macroevolutionary patterns of host use and butterfly diversity. Chapter I provides a mechanistic basis for such patterns through simulations of lineages evolved *in silico*. By separating the effects of the number of hosts used by a parasite lineage and the diversity of resources they encompass, we found that resource diversity, rather than host range *per se*, was the main driver of parasite species richness in both simulated and empirical systems. In Chapter II, we combined network and phylogenetic analyses to quantify support for the two main hypothesized drivers of diversification of herbivorous insects. Based on analyses of two butterfly families, Nymphalidae and Pieridae, we found that variability in host use is essential for diversification, while radiation following the colonization of a new host is rare but can produce high diversity. We then reconciled the two alternative hypotheses into a unified process of host-associated diversification where continuous probing of new hosts and retention of the ability to use hosts colonized in the past are the main factors shaping butterfly-plant networks. While network analysis is a powerful tool for investigating patterns of interaction, other methods are necessary to directly test the mechanisms generating the observed patterns. Therefore, in Chapter III we describe a model of host repertoire evolution we developed for Bayesian inference of evolution of host-parasite interactions. The approach was validated with both simulated and empirical data sets. Finally, in Chapter IV we used the method described in Chapter III to explicitly test the predictions made in Chapter II about the evolution of butterfly-plant networks. We found direct evidence for the role of expansion of fundamental host repertoire and phylogenetic conservatism as important drivers of host repertoire evolution. Thus, using three different approaches, we found overall support for the idea that variation in host use accumulated over evolutionary time is essential for butterfly diversification.

Keywords: *coevolution, host-parasite interaction, inference, networks.*

Stockholm 2019

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-163179>

ISBN 978-91-7797-568-7

ISBN 978-91-7797-569-4



Stockholm
University

Department of Zoology

Stockholm University, 106 91 Stockholm

EVOLUTION OF HOST REPERTOIRES AND THE DIVERSIFICATION
OF BUTTERFLIES

Mariana Pires Braga

Evolution of host repertoires and the diversification of butterflies

Mariana Pires Braga

©Mariana Pires Braga, Stockholm University 2019

ISBN print 978-91-7797-568-7

ISBN PDF 978-91-7797-569-4

Cover artwork by Rasmus Erlandsson

Photos by Niklas Janz

Printed in Sweden by Universitetsservice US-AB, Stockholm 2019

"If I see anything vital
around me, it is precisely
that spirit of adventure,
which seems
indestructible and is
akin to curiosity."

- Marie Curie

CONTENTS

Introduction.....	9
Modeling evolution in silico.....	11
Predicting patterns of interaction.....	14
Inference of host repertoire evolution.....	17
Conclusions	22
Svensk sammanfattning.....	23
References.....	24
Acknowledgements.....	26
List of chapters.....	27

INTRODUCTION

All organisms are ecologically specialized to some extent. Herbivorous insects have long served as models for the study of specialization (Futuyma and Moreno 1988) because of their impressive diversity and variation in plant resource use. This variation resulted in gradients of specialization among individuals, populations and species (Forister et al. 2015), and has been shown to mediate several ecological and evolutionary processes, such as the coexistence of competitors (Büchi and Vuilleumier 2014), persistence in face of environmental perturbation (Devictor et al. 2008), insect-plant network stability (Mougi and Kondoh 2012; Lever et al. 2014), and diversification (Janz et al. 2006; Hardy and Otto 2014).

Host specialization - leading to monophagy - only represents one end of the gradient, ignoring a great part of the variation in host range and in actual shifts between host plants (Janz et al. 2001; Agosta et al. 2010). Individuals of several herbivorous insects (e.g. most lepidopterans and hymenopterans) complete all of their larval development on an individual host, and are therefore monophagous individuals, or parasites (Thompson 1994). Other insects can move between and feed on two or more host individuals and are called grazers (Thompson 1994). Although grazers can feed on individuals of a same species (for example, in monocultures), grazing and polyphagy are generally correlated. Among parasitic insects, a variety of ecological, chemical, and genetic factors may cause host preferences to change locally, producing specialists with restricted diets in local communities (Fox and Morrow 1981). Variation in host use along spatial and temporal gradients can take place at the population level, when individuals are monophagous but feed on different plants, or at the species level, when monophagous populations are specialized on different hosts.

Due to the prevalence of host specificity and taxonomic conservatism among herbivorous insects, many studies have focused on the mechanisms and consequences of ecological specialization (Ehrlich and Raven 1964; Futuyma and Moreno 1988; Forister et al. 2012). This led to the generalization that host use by insects is both specialized and conservative (Ehrlich and

Raven 1964; Janz and Nylin 1998; Novotny and Basset 2005), which creates something of a paradox considering that insects have evidently colonized a good deal of the diversity of plants in a evolutionarily short time (Agosta et al. 2010). For instance, the ancestor of the butterflies probably colonized a relatively derived seed plant (Janz and Nylin 2008), hence the present patterns of host use are better explained by colonizations from ancestral plants onto the already diverse group of seed plants.

Even though a large-scale conservatism can be seen in most insect groups, rapid shifts and colonizations have been observed as a consequence of habitat change (Singer et al. 1993) and invasive or introduced plant species (Carroll and Boyd 1992; Fraser and Lawton 1994; Fox et al. 1997). In some cases, a complete shift in preferred host plant has occurred over a handful of generations (Singer et al. 1993), suggesting that under some circumstances adding new host plants might not be as difficult as one would expect. There is, therefore, strong evidence both at evolutionary and ecological time scales that host range expansions (i.e. increase in number of hosts used) are not rare.

There is less consensus, however, about the importance of host range expansions for diversification (Hamm and Fordyce 2015; Janz et al. 2016; Hamm and Fordyce 2016). Novel statistical approaches to investigate trait-dependent diversification have been developed recently (Maddison et al. 2007; FitzJohn 2012), but so far have produced divergent results when used to study host-associated diversification in butterflies (Hardy and Otto 2014; Hamm and Fordyce 2015; Hardy et al. 2016). Part of this problem may lie in the classification of host range into two opposing states (monophagous vs. polyphagous), which precedes most analytical methods. Strictly speaking, host range is not an independently evolving trait in its own right, but an emergent property of the underlying dynamics of gaining and losing host plants.

The ideal method to investigate the role of hosts in diversification processes should be able to deal with both the number of hosts used by each taxon and the identity of the plants in the repertoire. The trait that evolves across the insect evolutionary history is indeed the host repertoire, that is, the assemblage of hosts used by the insect, not just the number of hosts used (host range). Computational limitations still constrain the explicit modelling of host repertoire evolution in a statistical framework. Thus, in this thesis, I suggest different ways to tackle this problem, using a combination of methods to investigate the coevolutionary dynamics between parasites and their hosts, with a special focus on butterfly-plant interactions.

MODELING EVOLUTION IN SILICO

In Chapter I, we investigated the mechanistic basis for the origins of macro-evolutionary patterns of parasite diversity and host use, by simulating parasite lineages evolved *in silico*. We were particularly interested in how the colonization of new hosts affects the phenotypic distribution and the fitness landscape of parasites, and whether it intensifies diversification.

We described an individual-based model in which (i) parasites undergo sexual reproduction limited by genetic proximity, (ii) hosts are uniformly distributed along a one-dimensional resource gradient, and (iii) host use is determined by the interaction between the phenotype of the parasite and a heterogeneous fitness landscape. The model comprises three hierarchical levels: individual, species, and fitness landscape. A species is a group of individuals connected by reproduction, which only occurs between individuals with a genetic distance smaller than the mate recognition threshold. The fitness landscape is composed of fitness peaks, which represent hosts for the parasites (Fig 1). The position of a given host in the fitness landscape represents the optimum phenotype to use that given host (the parasite phenotype that yields maximum survival). Survival decreases with increasing difference between the host optimum and the parasite phenotype.

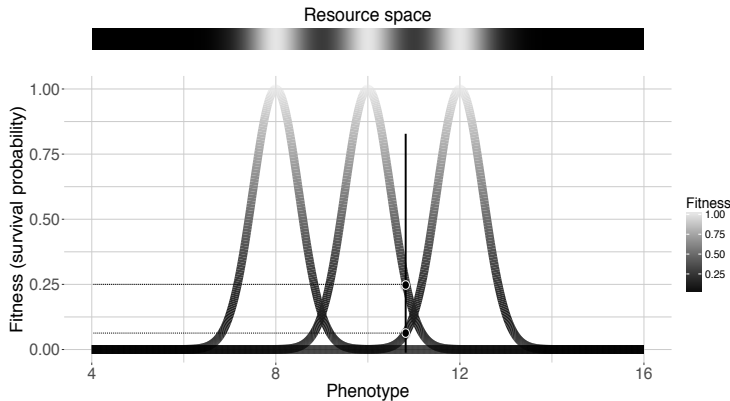


Figure 1 - Hypothetical fitness landscape with three hosts (peaks) and the resulting 1-dimensional resource space, where white patches represent phenotypes with positive fitness for each host. The solid vertical line is a projection of a given phenotype to show that the fitness of a parasite on a given host depends on where the line crosses the fitness curve (dots), and that this varies between hosts (this individual has a positive fitness in two hosts).

At each time step, three events happen in the following order: reproduction, host probing, and selection (Fig 2). The parasite population in a given host at a given time step is composed of the offspring of the parasites from the previous time step that did not attempt to colonize a new host and survived the selective pressure, in addition to the parasites that successfully colonized this given host. A complete description of model is given in Chapter I.

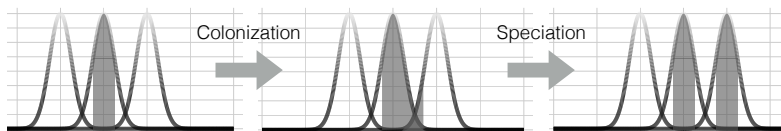


Figure 2 - Changes in the amplitude of the parasite phenotypic distribution (shaded area) after colonization of a new host (expansion to the fitness peak on the right), and after speciation (population splits in two).

Using this model, we performed simulations crossing parameter values, which resulted in 135 parameter combinations. Each combination was iterated for 1000 generations and replicated three times. At the end of each simulation, we recorded the total number of hosts used by all parasites (host range), the difference between maximum and minimum parasite phenotypes (phenotypic amplitude), and the number of isolated reproductive units (species richness). Then we estimated the effect of each model variable on parasite species richness using a Poisson regression, and on phenotypic amplitude and host range using Gaussian regressions. Finally, to estimate the relationships between model outcomes, we used partial correlations between species richness, phenotypic amplitude, and host range.

We found two main effects of host use on the evolution of parasite lineages. First, the colonization of a novel host allowed parasites to explore new areas of the resource space, increasing phenotypic and genotypic variation. Second, hosts produced heterogeneity in the parasite fitness landscape, which led to reproductive isolation and therefore, speciation. Parasite species diversity was maximized when hosts were at an intermediate distance in resource space, balancing the probabilities of colonization and divergent selection. Colonization happens quickly when hosts are very similar but divergent selection is stronger when hosts are distant.

In order to validate the results from our simulation study, we analyzed empirical data from butterfly-plant interactions. We used a metric of ordinated diet breadth (ODB) as a measure of phenotypic amplitude. This method uses information on how often different host plant taxa are utilized by the

same butterfly taxon as a proxy for resource similarity. We calculated ODB for the repertoire of host plant orders utilized by 43 butterfly tribes of Nymphalidae butterflies. We also calculated Faith’s phylogenetic diversity of plant orders used by each tribe (phylogenetic host range) and the taxonomic host range as the number of orders used by each tribe. We then used phylogenetic path analysis to assess the direct and indirect effects of these variables on species richness (number of species in the tribe).

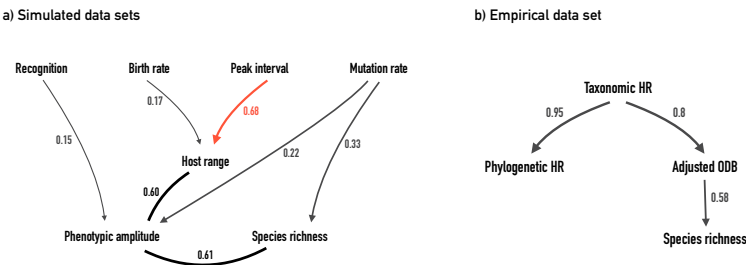


Figure 3 - Predictors of parasite species in simulated and empirical data sets. a) Relationships between model parameters and model outcomes. Arrows represent the effect of parameters on outcomes and associated numbers show the deviance explained by each parameter. Grey arrows show positive effects and the red arrow shows a negative effect. Lines connecting model outcomes indicate partial correlations between the three variables. b) Summary of phylogenetic path analysis for butterfly-plant interactions showing that the effect of taxonomic host range on parasite species diversity is mediated by host diversity (heterogeneity) in Nymphalidae tribes. Arrow thickness is scaled approximately with the standardized path coefficients, which are shown for each path.

On theoretical and empirical grounds, this study highlights the importance of the differentiation between host range and host diversity, with the latter having the main direct effect on diversification (Fig 3). Host range expansions lead to diversification, as long as they increase heterogeneity in the resource space, and consequently, in the fitness landscape. We see this as an important step forward in our understanding of diversity patterns generated by host-associated diversification.

In Chapter I we show that host use dynamics can drive diversification even in sympatry. In Chapter II we also investigate host-associated diversification, but focusing on the patterns of interaction generated by different processes.

PREDICTING PATTERNS OF INTERACTION

Evolutionary ecologists have long been interested in the causes and consequences of host shifts (Ehrlich and Raven 1964; Janz 2011). The various hypotheses of how colonization of new hosts leads to diversification can be placed along two main axes: (i) the relative prevalence of complete host shifts vs. expansion of the number of hosts, and (ii) the relative importance of key innovations vs. existing abilities (standing genetic variation and phenotypic plasticity) for colonization of new hosts. In Chapter II we compared two alternative extremes among these hypotheses, the adaptive radiation scenario and the variability scenario, which are also the two most prominent explanations for how changes in host use affect net diversification rates.

The adaptive radiation scenario hypothesizes that herbivorous insects quickly radiate into many species following a shift from an old to a novel plant taxon, by overcoming their defenses against herbivory. Therefore, host shifting (i.e., complete change in host use) is considered the main driver of diversification (Ehrlich and Raven 1964; Fordyce 2010). In contrast, the variability scenario predicts that diversification is maximized in insect taxa with the ability to use a wide range of potential hosts. The existence of such potential hosts – remnants of past host colonizations – makes host repertoires unstable over evolutionary time, since insects can mix and match between them relatively easily. The resulting oscillations in host range increase the chance of population fragmentation and thereby speciation, via both adaptive and neutral processes (Janz and Nylin 2008).

We first investigated whether signs of the above-mentioned scenarios of diversification can be seen in extant networks of interaction, by translating their predictions into network properties (Fig. 4). Radiations of herbivorous insects on new host plants should result in a modular network. Modularity emerges when the network contains recognizable subsets of taxa that interact more with each other than with other taxa in the network. Each radiation in a new host plant taxon should create a new module composed of closely related plants and insects, which descend from the ancestor that made the host shift. Alternatively, variability in host use should produce a nested network. Nestedness emerges if there is a specialist-generalist gradient in both trophic levels and the interacting assemblage of a taxon is a subset of the interacting assemblages of taxa with more interactions. According to the variability scenario, temporal variation in host range produces a specialist-generalist

gradient at any point in time, with ancestral hosts being used by both specialist and generalist insects, while novel/uncommon hosts are only used by generalists.

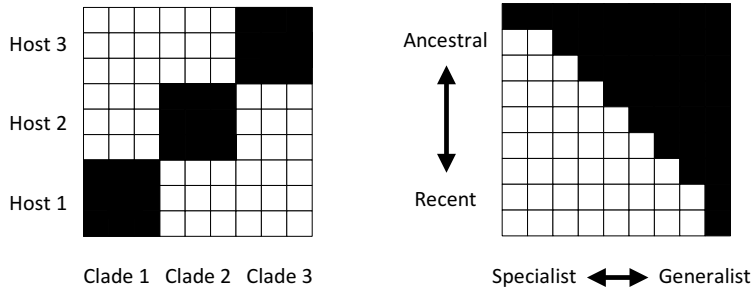


Figure 4 - Network structure predicted by the adaptive radiation (*left*) and the variability (*right*) scenarios. Both networks have insects in columns and hosts in rows. Black cells indicate insect-host interaction.

To validate our verbal arguments, we simulated insect diversification as taking place according to either the radiation or the variability scenario and then measured modularity and nestedness in the resulting networks. As expected, radiations produced modular networks while variability in host use produced nested networks.

The next step was to quantify the contribution of each scenario to empirical systems. We chose two butterfly families that were used as exemplars of each scenario: Pieridae, which is associated with adaptive radiations, and Nymphalidae, which is associated with the variability scenario. The Nymphalidae-plant network included 566 interactions between 295 Nymphalidae genera and 43 host-plant orders, and the Pieridae-plant network included 126 interactions between 67 Pieridae genera and 34 host-plant families. Despite the general acceptance that coevolution with host plants resulted in diversification in Nymphalidae and Pieridae in different ways, we found that the network structures of the two families are very similar. Both networks are nested and modular (Fig. 5), which indicates that the diversification of both families was influenced by both scenarios. By comparing the structure of the empirical networks to that of simulated networks, we found that the high levels of nestedness found in Pieridae and Nymphalidae suggest that the variability scenario played an important role on the diversification of both butterfly families. The modularity levels, however, could have emerged from radiations or

simply from phylogenetic conservatism in host repertoire, especially in Pieridae, where modularity is low.

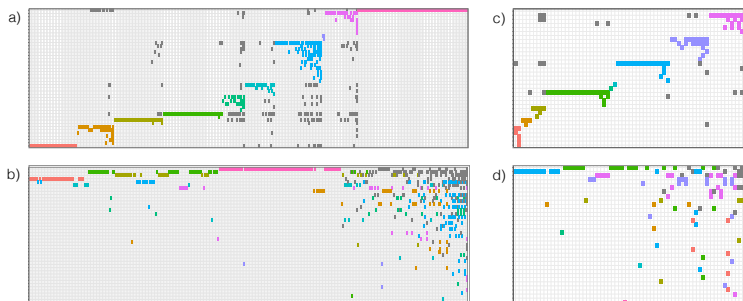


Figure 5 - Structure of butterfly-plant networks. (a-b) Nymphalidae genera in columns and host plant orders in rows. (c-d) Pieridae genera in columns and host plant families in rows. Colored cells are butterfly-plant interactions. Each color shows interactions within a module, and grey cells are interactions between modules. (a,c) Rows and columns sorted to emphasize modular affinity (order of modules is arbitrary). (b,d) Rows and columns sorted to emphasize nestedness, ordered from the upper right corner according to descending number of interactions.

We then performed additional analyses in order to assess the contributions of each scenario. First, we estimated the effect of each host plant on network structure and found that, as predicted, ancestral hosts produce network nestedness while novel hosts produce modularity. Finally, we measured the phylogenetic diversity within each module. In both networks, most modules were composed of closely related butterflies and distantly related plants.

Combining all results from Chapter II, we found that the modules in the studied networks are formed by grouping closely related butterflies that use a main host taxon (module hub). Several modules also include a number of other distantly related hosts that are used by a subset of the butterflies in the module, producing nestedness within modules. Hosts with a long coevolutionary history with the butterflies connect the various modules, producing overall network nestedness. These results led us to argue that the variability and radiation scenarios can be reconciled into a unified view of evolution of butterfly-plant interactions in which the continuous probing of new hosts allows both ongoing diversification through variability in host use and episodic radiations on new hosts.

INFERENCE OF HOST REPERTOIRE EVOLUTION

While Chapters I and II provide new insights into host-associated diversification using different approaches, the explicit modeling of host repertoire evolution has been hindered by methodological and computational constraints. Recent developments in phylogenetic Bayesian inference of evolution of discrete traits (Landis et al. 2013) considerably reduced these constraints, allowing the development of models of host repertoire evolution that can deal with the inherent complexity of the trait.

Inspired by the approach proposed by (Landis et al. 2013) for biogeographic inference, in Chapter III we describe a model of host repertoire evolution where parasite lineages can gain and lose hosts over time. In addition to allowing parasites to use a large number of host taxa simultaneously, this method also allows the modeling of host colonization as a two-step process, where first the parasite gains the ability to use a host and then starts to use it in nature. We implemented this feature by adding an intermediate state to the model, such that each host could assume one of three states for each parasite: 0 (non-host), 1 (potential host) or 2 (actual host). Thus, we could reconstruct not only realized host repertoires (composed by only actual hosts), but also fundamental host repertoires (composed of both potential and actual hosts). Another important feature of the model is the estimation of the effect of host phylogenetic relatedness on the likelihood of colonizing new hosts.

The model is implemented in RevBayes (Höhna et al. 2016), allowing us to perform simulation as well as Bayesian Markov chain Monte Carlo (MCMC) inference under the model. In Chapter III, we explored the statistical behavior of our model by simulating evolution of host-parasite interactions under a range of parameters. Overall, we were able to infer the true parameter values regardless of the level of phylogenetic conservatism in both parasites and hosts. We also compared the true coevolutionary history of each simulation to the corresponding posterior distribution of the sampled character histories. Estimation error was calculated as the sum of squared differences between estimated and true coevolutionary histories. Accuracy in the estimation of coevolutionary history was highest when the degree of phylogenetic conservatism on both butterflies and plants was high. Overall, error was higher on the estimation of actual hosts (state 2) than potential hosts (state 1), but both were within acceptable levels.

We then demonstrated the empirical application of our approach with a Bayesian inference of the coevolutionary history between 34 Nymphalini butterflies and 25 angiosperm families (Fig 6). We estimated the rate of host repertoire evolution along the Nymphalini tree as being between 0.53 and 1.22 events per million years. Bayes factors favored the independence model, where the probability of gaining a given hosts is not affected by the phylogenetic distance between hosts.

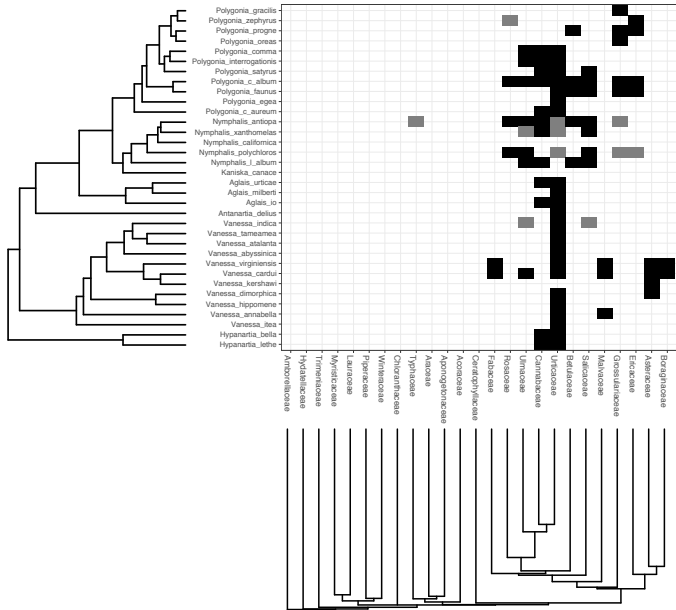


Figure 6 - Data set of interactions between Nymphalini butterflies (rows) and angiosperm host plants (columns). Black cells show actual hosts (interaction recorded in nature) and grey cells show potential hosts (larvae were able to feed on host during establishment tests).

Gain and loss rate estimates were not symmetric, and that also varied between states. The gain of the ability to use a host was estimated to be very rare (0.25% to 1% of overall rate), whereas loss was common (34% to 78% of overall rate). Transition rates between states 1 and 2 were more symmetric and gain was more common than loss (1→2 between 17% and 54%; 2→1 between 3% and 15% of overall rate). These results support the idea that re-colonization of hosts that were used by ancestors of extant butterflies might be more common than independent colonizations of the same host taxon by phylogenetically widespread butterflies (Janz et al. 2001).

As the approach described in Chapter III was able to accurately infer co-evolutionary histories in data sets simulated under a range of parameter combinations, we used it to cross-validate the network approach used in Chapter II and test the main conclusions from that study. For that, we reconstructed the historical interactions between Pieridae butterflies and their host plants (same data set as used in Chapter II). We modeled host repertoire evolution along a phylogenetic tree containing 66 genera of Pieridae. For the host tree, we pruned the phylogenetic tree for angiosperm families by keeping the 33 families known to be hosts of pierid butterflies and then collapsing the remaining branches to more ancestral nodes until only 50 terminal branches were left. Host repertoires were, therefore, composed of 50 hosts. By doing this pruning, we ensured that all angiosperm lineages were represented in the analysis, at the same time as keeping the number of hosts within the limits of computational tractability.

One of the main ideas we tested was that the evolution of butterfly-plant networks is mainly driven by the probing of potential hosts combined with phylogenetic conservatism in host-use abilities. We estimated the rate of repertoire evolution along the Pieridae tree as being between 0.055 and 0.11 events per million years, which is ten times slower than the estimated rate for Nymphalini butterflies. The asymmetry between transition rates, however, was very similar to Nymphalini. The gain $0 \rightarrow 1$ was estimated to be very rare (0.6% of overall rate), whereas the loss $1 \rightarrow 0$ was common (66%). Conversely, transition between states 1 and 2 were similar (14% and 19% of overall rate), which means that recolonization of a host that was used in the past is easier than the colonization of a completely new host. Together, these results provide the basis for the observed phylogenetic conservatism in the host repertoires of pierid butterflies.

We then reconstructed ancestral networks at two time slices during the diversification of Pieridae (Fig. 7). We selected seven internal nodes in the Pieridae tree between 60 and 50 Mya to reconstruct the first ancestral network. For the second ancestral network, we selected 16 nodes between 40 and 20 Mya. The two ancestral networks were composed of interactions between each selected node and the hosts that were in states 1 or 2 in at least 95% of the sampled histories. This means that only interactions with more than 0.95 posterior probability were included. Figure 8 shows the posterior probability of every possible interaction in the ancestral network at 40-20 Mya.

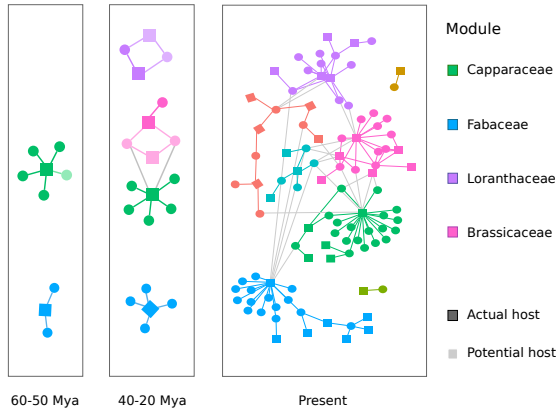


Figure 7 - Two ancestral networks and extant network of interactions between Pieridae and Angiosperms. Butterflies are represented by circles and plants by squares. Colors show network modules.

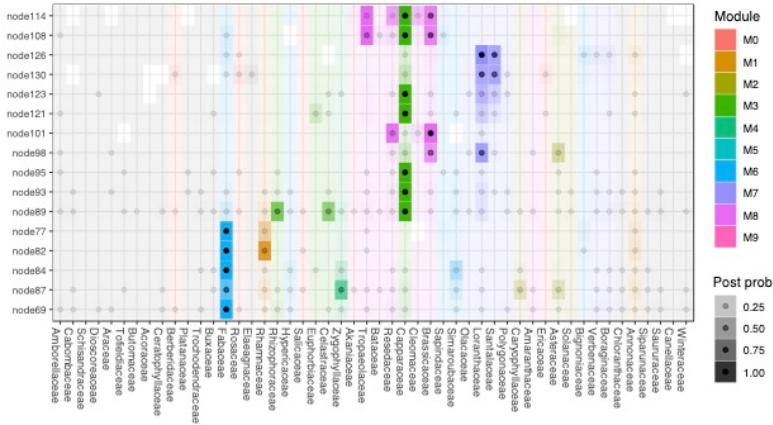


Figure 8 - Posterior probabilities for interactions between ancestors of extant pierid butterflies (internal nodes) and angiosperm families at 40-20 Mya. Squares are colored according to the module that the host belongs to in the extant network (see Figure 5c). Black dots mark the interactions with actual hosts (state 2). Opacity indicates interaction probability. Note that almost all interactions have been sampled at least once.

In Chapter II we suggest that the evolution of butterfly-plant interactions goes through three phases of change in network structure. The reconstructed evolutionary history of interactions between Pieridae and their host plants presented in Chapter IV exemplifies each one of them. The first ancestral network (60-50 Mya) shows the origin of a new module (green) when early

Pierinae switched to Capparaceae. At least until 50 Mya, those butterflies were specialized on the new host group, resulting in a network composed of two independent modules of closely related butterflies. Thus, this network exemplifies the first phase in the evolution of butterfly-plant interactions.

In the second ancestral network (40-20 Mya), Pierinae butterflies started to expand their host repertoires, adding new potential and actual hosts. This expansion promoted the formation of new modules (pink and purple), but the retention of ancestral hosts (in this case, Capparaceae) kept these modules connected. Thus, this network exemplifies the second phase, when probing of new hosts increases network connectance. Interestingly, this network represents the origin of the four main modules in the extant network, which are formed by butterflies associated to the module hubs identified in Chapter II (Fig. 9).

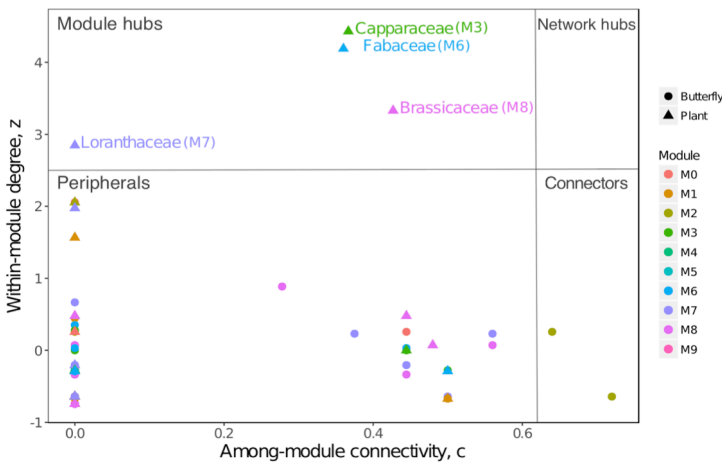


Figure 9 - Relative contribution of each taxon to the modular structure, defined by measures of connectivity with taxa within the same module (y-axis) and with taxa assigned to other modules (x-axis) of the network. Names of modules and network hubs are shown. Colors of modules are consistent with Figs. 7 and 8.

Finally, the third phase, which is characterized by network nestedness, is exemplified by the network at present time, where the number of both butterflies and hosts increased in each module, as well as the number of connections between modules. The colonization of new hosts combined with the recolonization of ancestral hosts (mainly Fabaceae and Capparaceae) resulted in the observed nestedness in the extant network.

CONCLUSIONS

Understanding how ecological interactions change is crucial to explain phenomena at various timescales, such as the emergence of infectious diseases, community assembly, and parasite diversification (Hoberg and Brooks 2015). By analyzing the same data set with completely independent approaches (Chapters II and IV), we validated the use of network analysis to test hypotheses about the evolution of ecological interactions and provided direct support for the unification proposed in Chapter II to explain the diversification of herbivorous insects.

The approach we describe in Chapter III was designed to reconstruct and quantify changes in host-parasite associations by modeling the process of gaining and losing hosts. Thus, it allowed us to explicitly model features that are intrinsic to the system, such as the existence of potential hosts and the effect of host phylogenetic relatedness on the colonization of new hosts. In this thesis, this approach was used to investigate host-associated diversification in butterflies, but it has potential to address various questions in evolutionary ecology. The implementation of the model in RevBayes facilitates future development to incorporate more of the complexity in species interactions.

As a whole, this thesis contains strong evidence for the role of host range expansions on the evolution of butterfly-plant interactions and on butterfly diversification. Even though closely related butterflies tend to use similar repertoires of host plants, that does not prevent the probing of new hosts. Part of the observed phylogenetic conservatism in host repertoires comes from the retention of the ability to use hosts used in the past, which facilitates recolonization. Such ancestral abilities favors changes in host repertoires, which increase the diversity of hosts used by a clade. The resulting host diversity, in turn, promotes butterfly diversification.

SVENSK SAMMANFATTNING

Alla växtätande insekter är i någon mån specialiserade på sina värdväxter, men specialiseringens omfattning är olika beroende på art. Samevolution mellan insekter (parasiter) och växter (värdorganismer) åberopas ofta som förklaringen till den stora artmångfald som uppvisas bland insekter, men det råder delade meningar om hur pass stor inverkan specialisering faktiskt har på artbildning. Trots att ämnet har utforskats under lång tid förbättras fortfarande vår förståelse av evolutionen bakom samspelet mellan värdväxter och parasiter, och nyare teoribildning understryker vikten av breddning (att kunna tillgodogöra sig nya värdväxter) för artbildning hos parasiter. I den här avhandlingen använder vi oss av olika perspektiv för att i detalj undersöka ursprunget till makroevolutionära mönster mellan värdväxtutnyttjande och artbildning hos dagfjärilar.

Kapitel I avhandlar, genom datorsimuleringar, mekanistiskt de processer som leder fram till de makroevolutionära mönstren. Genom att åtskilja effekten av *antal* ätbara värdväxter hos en grupp fjärilar från hur *olika* värdväxterna är sinsemellan kunde vi visa att olikhet, snarare än antal i sig, ligger bakom artmångfald hos parasiter, både i simulerade och verkliga system.

I kapitel II kombinerade vi nätverks- och släktskapsanalyser för att jämföra stödet för de två olika drivkrafter som föreslagits som huvudorsaker till artbildning inom växtätande insekter. Utifrån två dagfjärilsfamiljer, Nymphalidae och Pieridae, drog vi slutsatsen att bredd i växtutnyttjande är avgörande för artbildning. Men det visade sig också att när insekter klarar att tillgodogöra sig en ny värdväxt så kan det ge upphov till snabb artbildning, även om det är förhållandevis ovanligt att det inträffar. Vi kunde på så vis förena de två olika föreslagna drivkrafterna och visa att artbildning kopplad till växtutnyttjande gynnas både av koloniseringar av nya värdväxter och av ett kontinuerligt utforskande av de växter som använts historiskt, och som de alltså inte helt förlorat förmågan att äta.

Trots att nätverksanalys är en utmärkt metod för att kartlägga mönster i samspel mellan arter så behövs det andra metoder för att uttryckligen testa vilka mekanismer som ger upphov till mönstren ifråga. I kapitel III utvecklade vi därför en evolutionär modell för värdväxtutnyttjande som vi kunde analysera med hjälp av bayesianska metoder. Tillvägagångssättet utvärderades med både simulerade och empiriska data.

I kapitel IV använde vi oss slutligen av metoden ifrån kapitel III för att uttryckligen testa förutsägelseerna om evolution av fjärl-växt-nätverket i kapitel II. Vi hittade stöd för att den viktigaste drivkraften är ett kontinuerligt utforskande av den repertoar av potentiella värdväxter som de ärvt från sina förfäder. Därmed kunde vi, med hjälp av tre olika metoder, sluta oss till att skillnader i värdväxtutnyttjande som uppstått under evolutionära tidsrymder, är avgörande för artbildning bland fjärilar.

REFERENCES

- Agosta S.J., Janz N., Brooks D.R. 2010. How specialists can be generalists: resolving the “parasite paradox” and implications for emerging infectious disease. *Zoologia*. 27:151–162.
- Büchi L., Vuilleumier S. 2014. Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. *The American Naturalist*. 183:612–624.
- Carroll S.P., Boyd C. 1992. Host Race Radiation in the Soapberry Bug - Natural-History with the History. *Evolution*. 46:1052–1069.
- Devictor V., Julliard R., Jiguet F. 2008. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos*. 117:507–514.
- Ehrlich P.R., Raven P.H. 1964. Butterflies and plants: a study in coevolution. *Evolution*. 18:586.
- FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*. 3:1084–1092.
- Fordyce J.A. 2010. Host shifts and evolutionary radiations of butterflies. *Proceedings of the Royal Society B: Biological Sciences*. 277:3735–3743.
- Forister M.L., Dyer L.A., Singer M.S., Stireman J.O.I., Lill J.T. 2012. Revisiting the evolution of ecological specialization, with emphasis on insect-plant interactions. *Ecology*. 93:981–991.
- Forister M.L., Novotny V., Panorska A.K., Baje L., Basset Y., Butterill P.T., Cizek L., Coley P.D., Dem F., Diniz I.R., Drozd P., Fox M., Glassmire A.E., Hazen R., Hrcek J., Jahner J.P., Kaman O., Kozubowski T.J., Kursar T.A., Lewis O.T., Lill J., Marquis R.J., Miller S.E., Morais H.C., Murakami M., Nickel H., Pardikes N.A., Ricklefs R.E., Singer M.S., Smilanich A.M., Stireman J.O., Villamarin-Cortez S., Vodka S., Volf M., Wagner D.L., Walla T., Weiblen G.D., Dyer L.A. 2015. The global distribution of diet breadth in insect herbivores. *Proc Natl Acad Sci USA*. 112:442–447.
- Fox C.W., Nilsson J.A., Mousseau T.A. 1997. The ecology of diet expansion in a seed-feeding beetle: Pre-existing variation, rapid adaptation and maternal effects? *Evol Ecol*. 11:183–194.
- Fox L.R., Morrow P.A. 1981. Specialization: species property or local phenomenon? *Science*. 211:887–893.
- Fraser S.M., Lawton J.H. 1994. Host-Range Expansion by British Moths Onto Introduced Conifers. 19:127–137.
- Futuyma D.J., Moreno G. 1988. The Evolution of Ecological Specialization. *Annual Review of Ecology and Systematics*. 19:207–233.
- Hamm C.A., Fordyce J.A. 2015. Patterns of host plant utilization and diversification in the brush-footed butterflies. *Evolution*. 69:589–601.
- Hamm C.A., Fordyce J.A. 2016. Greater host breadth still not associated with

- increased diversification rate in the Nymphalidae—A response to Janz et al. *Evolution*. 70:1156–1160.
- Hardy N.B., Otto S.P. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society B: Biological Sciences*. 281:20132960–20132960.
- Hardy N.B., Peterson D.A., Normark B.B. 2016. Nonadaptive radiation: Pervasive diet specialization by drift in scale insects? *Evolution*. 70:2421–2428.
- Hoberg E.P., Brooks D.R. 2015. Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 370:–20130553.
- Höhna S., Landis M.J., Heath T.A., Boussau B., Lartillot N., Moore B.R., Huelsenbeck J.P., Ronquist F. 2016. RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Systematic Biology*. 65:726–736.
- Janz N. 2011. Ehrlich and Raven Revisited: Mechanisms Underlying Codiversification of Plants and Enemies. *Annu. Rev. Ecol. Evol. Syst.* 42:71–89.
- Janz N., Braga M.P., Wahlberg N., Nylin S. 2016. On oscillations and flutterings—A reply to Hamm and Fordyce. *Evolution*. 70:1150–1155.
- Janz N., Nyblom K., Nylin S. 2001. Evolutionary dynamics of host-plant specialization: a case study of the tribe Nymphalini. *Evolution*. 55:783–796.
- Janz N., Nylin S. 1998. Butterflies and plants: a phylogenetic study. *Evolution*. 52:486.
- Janz N., Nylin S. 2008. The oscillation hypothesis of host-plant range and speciation. In: Tilmon K.J., editor. *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects*. Univ. California Press p. 203–215.
- Janz N., Nylin S., Wahlberg N. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evol Biol.* 6:4–10.
- Landis M.J., Matzke N.J., Moore B.R., Huelsenbeck J.P. 2013. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*. 62:789–804.
- Lever J.J., van Nes E.H., Scheffer M., Bascompte J. 2014. The sudden collapse of pollinator communities. *Ecology Letters*. 17:350–359.
- Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a Binary Character's Effect on Speciation and Extinction. *Systematic Biology*. 56:701–710.
- Mougi A., Kondoh M. 2012. Diversity of interaction types and ecological community stability. *Science*. 337:349–351.
- Novotny V., Basset Y. 2005. Host specificity of insect herbivores in tropical forests. *Proceedings of the Royal Society B: Biological Sciences*. 272:1083–1090.
- Singer M.C., Thomas C.D., Parmesan C. 1993. Rapid Human-Induced Evolution of Insect Host Associations. *Nature*. 366:681–683.
- Thompson J.N. 1994. *The Coevolutionary Process*. University of Chicago Press.

ACKNOWLEDGEMENTS

It has been more than eight years since I stepped into Walter Boeger's lab at Federal University of Paraná in Brazil and discovered the world of host-parasite interactions. In the course of these years I was lucky enough to work with amazing people, from whom I learned about everything I know today. Some of the ideas in this thesis actually come from spontaneous meetings with Walter, when I would just go to him with my head full of questions and we would talk (and draw!) until we found some resolution. I really don't know what I would be doing today if I hadn't worked with you, Walter. Thank you so much for everything!

Even though I had a great supervisor, I wanted to go abroad for my PhD. Choosing to move to Stockholm to work with Niklas and Sören was the easiest and best decision I made. From the first emails, I knew I was in good hands. And over these four years, my admiration for you two just increased. Thank you for the freedom you gave me to try new things, the trust you always had on me, and the constant support during these intense years. I couldn't ask for better supervisors.

I was also very lucky to have Fredrik as a co-supervisor. I learned so much from you! Thank you for finding the time for me, and for your enthusiasm with every progress we made. It was also great to meet all the board-game lovers at Naturhistoriska riksmuseet and be part of the "Fredrikers".

There are many current and former members of Zootis that I wish to thank. Kalle and Chris, thank you for being a great follow-up group, being my reality-check and making sure that I kept on track. Siw, Minna, Anette and Ulf, you guys are the best. Thank you so much for making life that much easier for us. As for the Zootis Party People (specially the original ones), you have been my extended family over these four years. This department is truly special, and although I could name the people that have been my closest friends, I do believe that everyone contributes to the special friendly atmosphere at Zootis.

One person, however, has been extra special. Thank you, Rasmus, for all the random conversations along these years about laundry, recycling or public housing programs in Sweden. Thank you for being by my side during these busy months. It really made a big difference!

None of this would have been possible without the support of my family. My parents always encouraged me to go after what I wanted, even if it meant to be apart from them. It was not easy to move so far away, but Sweden received me really well (even though it was a rainy November). It has been a pleasure to live and work in this country. I hope to come back soon!

LIST OF CHAPTERS

- I. **Braga, MP**, SBL Araujo, S Agosta, D Brooks, E Hoberg, S Nylin, N Janz and WA Boeger. (2018) Host use dynamics in a heterogeneous fitness landscape generates oscillations in host range and diversification. *Evolution*, vol. 72(9), pp. 1773–1783.
- II. **Braga, MP**, PR Guimarães Jr., CW Wheat, S Nylin and N Janz. (2018) Unifying host-associated diversification processes using butterfly-plant networks. *Nature Communications*, 10.1038/s41467-018-07677-x.
- III. **Braga, MP**, M Landis, S Nylin, N Janz, F Ronquist. Bayesian analysis of host repertoire evolution. - *Manuscript*.
- IV. **Braga, MP**, M Landis, S Nylin, N Janz, F Ronquist. Evolution of butterfly-plant networks revealed by Bayesian inference of host repertoire. - *Manuscript*.

Candidate contributions to thesis articles*

	I	II	III	IV
Conceived the study	Substantial	Substantial	Significant	Substantial
Designed the study	Significant	Substantial	Substantial	Substantial
Collected the data	Substantial	Minor	Substantial	Minor
Analysed the data	Substantial	Substantial	Substantial	Substantial
Manuscript preparation	Substantial	Substantial	Substantial	Substantial

* Contribution Explanation

Minor: contributed in some way, but contribution was limited.

Significant: provided a significant contribution to the work.

Substantial: took the lead role and performed the majority of the work.