

Diversity of plant-feeding insects: using informatics to explore the effects of natural enemy interactions

by

Mayrolin García Morales

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 7, 2016

Keywords: biodiversity, speciation, extinction, macroevolution

Approved by

Nate B. Hardy, Assistant Professor, Department of Entomology and Plant Pathology
Alana L. Jacobson, Assistant Professor, Department of Entomology and Plant Pathology
Akito Y. Kawahara, Assistant Professor / Curator, Florida Museum of Natural History

Abstract

The aim of my thesis research was to gain a deeper understanding of the species diversity of plant-feeding insects. I pursued that aim with two complementary projects. In the first, I developed a web application to manage information about the species diversity of scale insects. In the second, I used scale insect data, and leafroller moth data, to test if the push from natural enemies and the refuge of new host plants is what drives ecological change and speciation in plant-feeding insects.

Acknowledgements

This work is dedicated to the star of the lab, the crown of thorns, a cutting from a *Euphorbia milii* specimen belonging to Dr. Kathy Flanders. Thank you for surviving two years of neglect. It was an inspiration.



I would like to thank the people who made this work possible, in no particular order of importance. I would like to thank my family. This work would be meaningless without you. I want to thank the rest of the Hardy Lab, Ricki Hamilton and Nick Christodoulides, for their support and encouragement. Special thanks go to my committee members, Dr. Alana Jacobson and Dr. Akito Kawahara, for their help and guidance. I would like to thank everyone in the Entomology and Plant Pathology department for making my time in Auburn unforgettable. Lastly, I would like to thank my major advisor, Dr. Nate Hardy, whose guidance, mentorship, and life advice were indispensable to my success. Thank you for making me a better scientist.

Table of Contents

Abstract	ii
Acknowledgments	iii
List of Tables	v
List of Illustrations	vi
List of Abbreviations	vii
Chapter 1: Introduction	1
Chapter 2: Rebooting ScaleNet	6
Background	6
Methods	8
Results and Discussion	11
Chapter 3: Natural enemy pressure drives diversification in plant-feeding insects	17
Background	17
Methods	19
Results	25
Discussion	30
Chapter 4: Conclusions	33
References	35

List of Tables

Table 1: Results from tri-trophic and bi-trophic generalized linear models	27
Table 2: Results from bitrophic QuaSSE models	29

List of Figures or Illustrations

Figure 1: Diagram of the new ScaleNet data model	10
Figure 2: Annotated phylogeny of scale insects	24
Figure 3: Annotated phylogeny of leafroller moths	25
Figure 4: Scatterplot for tritrophic models for scale insects and leafroller moths	28

List of Abbreviations

AIC	Akaike's Information Criterion
BASIS	Biological and Systematic Information System
BS	bootstrap
dAIC	AIC difference to the distribution of AIC differences calculated in simulations
HR	host range
ML	maximum likelihood
MVC	model-view-controller
NE	natural enemy
QuaSSE	Quantitative State Speciation and Extinction
SSE	state dependent speciation and extinction

Chapter 1: Introduction

Overview

There are close to 1.9 million described species of eukaryotic organisms (1). About 1 million are species of insects (2). By all estimates we have described only a fraction of the true extant species diversity (3, 4). How many species are living on Earth today? Why are so many of them insect species? What evolutionary processes drive the species diversification of insects? How do these processes differ from those in other, non-insect, evolutionary lineages?

Biologists since Darwin have been asking these questions, but until a few years ago we were not able to address them with any rigor. Advances in computing power and in the sharing of information have made it possible to assemble much more comprehensive and accessible catalogues of biodiversity. Computational and theoretical advancements have also given us the means of using the data in biodiversity catalogues to test hypotheses about the macroevolutionary history of insects.

My thesis work was motivated by a desire to a) learn how to develop and manage information about species diversity, and b) to understand the processes that promote species diversity. First, I used open-source tools to rebuild ScaleNet, an online database of scale insect biodiversity. ScaleNet is the definitive source of information for scale insects (Hemiptera: Coccoidea), and is used heavily by many biologists. At the time I was starting my MS work, the ScaleNet software platform was crumbling and, short of a major re-design and reimplementation, was going to be taken offline. My work on ScaleNet, described in Chapter 2, has been published in the journal *Database* (5). I then used data for scale insects, along with

data I was able to glean for leafroller moths (Lepidoptera: Tortricidae), to investigate whether natural enemy interactions drive ecological and species divergence in plant-feeding insects. Our results suggest they do by showing that natural enemy pressure is positively correlated with species richness and speciation rates in these two plant-feeding groups.

Through my thesis research, I have arrived at a much richer understanding of the history of taxonomic diversification in plant-feeding insects, and have gained facility with the informatics tools and analytical methods that biologists use to investigate hypotheses about insect biodiversity. I have also been able to significantly improve the plant-feeding insect biodiversity infrastructure, and advance the macroevolutionary theory.

Theoretical Background

Most insect species are herbivores. Most of our ideas for why plant-feeding insects are so species rich fixate on their trophic biology, that is, plant eating. Fifty-two years ago, Ehrlich and Raven (1964) proposed the Escape and Radiate Hypothesis, which explains the diversity of plant-feeding insects and their host plants as the result of co-evolutionary adaptive radiation (6): a plant lineage that evolves a novel chemical defense escapes its herbivores. This relaxes stabilizing selection, loosens the constraints on plant evolution, and increases the likelihood of ecological diversification and speciation. Likewise, an insect lineage that evolves a counter adaptation to a novel plant defensive chemistry escapes from competition with other plant-feeding insects for limited plant resources. This also unties constraints on evolution and promotes diversification. The macroevolutionary history of plants and their insect herbivores is a

back and forth between defensive innovation and circumvention, with alternating phases of repressive antagonism and unfettered diversification.

The Escape and Radiate Hypothesis motivated decades of research into the function of plant secondary chemistry, and in the minds of generations of evolutionary biologists, it ensconced plant-insect coevolution as the main factor driving plant-feeding insect speciation (7). Since 1964, biologist's explanations for why plant-feeding insects are so diverse have mainly been elaborations of the Escape and Radiate Hypothesis even though there is no support for all of the underlying assumptions of this hypothesis for ecological change and speciation.

Various hypotheses have been proposed since 1964 to address limitations to the Escape and Radiate hypothesis. The Escape and Radiate Hypothesis does not tell us *why* relaxed natural selection should result in many species. Undoing constraints on the evolution of a plant-feeding insect lineage could just make that lineage more ecologically generalized, for example by allowing it to feed on more host plants. Why should the evolution of a counter-adaptation to a plant defense result in many host-specific species instead of a few generalist species? To address this Janz and Nylin's (2008) proposed the Oscillation Hypothesis (8). The diet breadth of a group of plant-feeding insects balloons in response to a key evolutionary innovation that lets them feed on plants that had been protected by a chemical defense. Then performance tradeoffs across host plants select against large diets, and drives the evolution of host-plant specialist species. Tradeoffs in host use – in which genotypes that do well on one host do poorly on another -- are essential to the Oscillation Hypothesis, but there is little evidence for these tradeoffs (7). How could host-use evolution drive plant-feeding insect

speciation if there are no tradeoffs that select against broad diets? Hardy and Otto (2014) proposed the Musical Chairs Hypothesis: Speciation is decoupled from host specialization, and instead is driven by host-plant switching in generalists and specialists alike (9). Further elaborations of the Escape and Radiate Hypothesis have been hampered by the fact that until a few years ago, we lacked the wherewithal to test its macroevolutionary predictions. For much of the past fifty years, the theory of plant-feeding insect diversification hindered by the absence of appropriate datasets and analytical frameworks for testing these hypotheses. They have also lacked the integration of other species interactions that are playing a role in diversification and speciation.

Plant-feeding insects have intimate relationships with their hosts. A host plant is something to eat, a home, a place to meet mates, and a prize to fight over. It makes sense to start with the host plants when thinking about the ecology of plant-feeding insects, however, interactions with other trophic levels, such as natural enemies, may also influence diversification and speciation. In fact, the reason that biological control is effective pest management is that natural enemies can kill most of their prey (10). Specialist natural enemies in particular can decimate populations of plant-feeding insects, with parasitoids alone accounting for up to 65% of mortality in some studies (11). Plant-feeding insects consider the risk of natural enemy attack when making decisions about host use (12-15), and will choose host-plants that have a smaller risk of predation even if it comes at a cost of nutrition (12-15).

The tri-trophic niche concept of Singer and Stireman (2005) explicitly considers the effects of natural enemy interactions over evolutionary time: Population persistence depends on

the interaction between host-plant associated physiological performance and mortality from natural enemies. Plant-feeding insects experience destabilizing selection from their natural enemies, and one way that they can respond to this selection is by evolving host-plant use. Evolutionary switches in host-plant use are only favored when a novel host provides a better combination of nutrition and enemy-free space. Singer and Stireman (2005) predict that diversity should be highest in lineages that experience high divergent selection from natural enemy pressure and have high ecological opportunity to escape that pressure, by evolving host use. Here we test these predictions using comparative phylogenetic tools. Is natural enemy pressure positively correlated with species richness and speciation rates? Have we been focusing too exclusively on the evolutionary relationships between plant-feeding insects and their hosts?

Chapter 2: Rebooting ScaleNet

Introduction

Scale insects are sap-sucking plant parasites that can be found almost anywhere that plants grow. They get their name from the protective waxy exudates produced by most species. Currently, there are at least 8194 described species, classified among 50 families. Scale insects play key roles in ecosystems. They, along with most other plant-feeding members of the order Hemiptera, are the only insects that feed exclusively on phloem sap (although armored scale insects feed primarily on parenchyma cells) (16). Phloem is rich in sugars but poor in amino acids, and phloem-feeding is an inefficient process. The waste is copious amounts of honeydew, i.e. sugar-rich excrement that is an important food source for birds, mammals and especially other insects (16). The availability of honeydew can affect insect communities in ways that alter ecosystem processes such as herbivore assemblage, soil structure, and predation (17, 18). Many scale insect species are agricultural pests, damaging plants through sap loss, encouraging the growth of sooty molds and vectoring plant diseases. Scale insects can be difficult to detect, and are extremely invasive. For example, scale insects account for only 1% of the total insect fauna of the United States, but for 13% of the introduced insect fauna, and on average one new invasive species is established as a pest in the USA per year (19). The host plant associations of scale insects have been exceptionally well documented, and the breadth of these associations is unusually variable. As is the case for other plant-feeding insects, most scale insect species are host-plant specialists. However, some species are among the most

polyphagous insect species known. For example, the brown soft scale, *Coccus hesperidum*, can successfully develop on host plants in at least 121 plant families, and 325 plant genera. Scale insects are also noteworthy for the unparalleled diversity of their genetic systems, and for the diversity and complexity of their relationships with endosymbionts (20). In addition to being a taxing problem for applied biologists, they are emerging as models for research addressing questions about the evolution of reproductive modes, genetic conflict and collaboration, and niche breadth evolution. There is high demand for synoptic information about the biological diversity of scale insects. That demand is met by ScaleNet.

ScaleNet

ScaleNet is a manually-curated, web-accessible database that models the biological diversity of scale insects through 300 years of published research. ScaleNet manages information about the systematics, ecological associations (host plants, natural enemies and mutualists), geographic distributions, life histories, economic importance and morphology of each scale insect species. As a model of the scale insect literature, the core of ScaleNet is an exhaustive bibliography. The rest of the information in the database can be thought of as annotations of that literature. ScaleNet began as a collaboration between Yair Ben-Dov (Agricultural Research Organization, Israel Department of Entomology), Douglass R. Miller (US Department of Agriculture) and Gary A.P. Gibson (Agriculture and Agri-Food Canada), with funding from the USA–Israel Binational Agricultural Research and Development Fund. It was developed as a Microsoft FoxPro application, using the BASIS (Biological and Systematic

Information System) database schema engineered by Gary Gibson and Jennifer Read (Agriculture and Agri-Food Canada) to manage taxonomic bibliographies. It first went online in 1995 (21). For 20 years, the ScaleNet data grew and evolved, but the ScaleNet application did not. By 2015, ScaleNet was running on an unsupported, insecure, closed-source software platform and was no longer tenable. Here, we describe a new version of ScaleNet.

Methods

Redeveloping ScaleNet

Our overarching goals for the redevelopment of ScaleNet were to (i) keep it online, (ii) make the software and data store easier to maintain, (iii) improve quality control, and (iv) make it easier to extend and articulate with other biodiversity resources. Our new version of ScaleNet is a Django application (a Python web framework: <https://www.djangoproject.com/>) with an SQLite database engine (<https://www.sqlite.org/>), that currently runs on Linux, behind an Apache web server (<http://httpd.apache.org/>), but which can be configured to run in other environments. Django follows a Model-View-Controller architecture, i.e. the controller (logic) receives user requests and fetches information from the model (data store) to be displayed in a view (HTML). We normalized the data model (Figure 1) and performed the data migrations with a set of custom Python scripts. As part of the migration we performed a number of data cleaning and standardization routines. We standardized the valid scientific names and classifications of all ecological associates following the schema of the Catalogue of Life (CoL:

<http://www.catalogueoflife.org/>) 2015 annual checklist. To amend spelling errors in the names of ecological associates, we used the fuzzy matching feature of the Global Names Resolver API (<http://resolver.globalnames.biodinfo.org/>). In addition to adding the CoL classification of ecological associates to the ScaleNet schema, we added a class for the classification of scale insect taxa (absent from the original schema) and another for nested relationships of the geopolitical and zoogeographical units that are used to describe the geographic distributions of scale insects. As ScaleNet is a model of the scale insect primary literature, all ScaleNet data need to be associated with a publication. However, early in its initial development, ScaleNet was seeded with information from databases compiled by Y.B-D. to summarize the biological diversity of the scale insect families Coccidae and Pseudococcidae (22,23). At that time validation sources for host and distribution records were not being recorded. These data are invalid in the new ScaleNet, and were not migrated. Instead, they were flagged and given to the ScaleNet curators to be manually restructured and added to the new database.

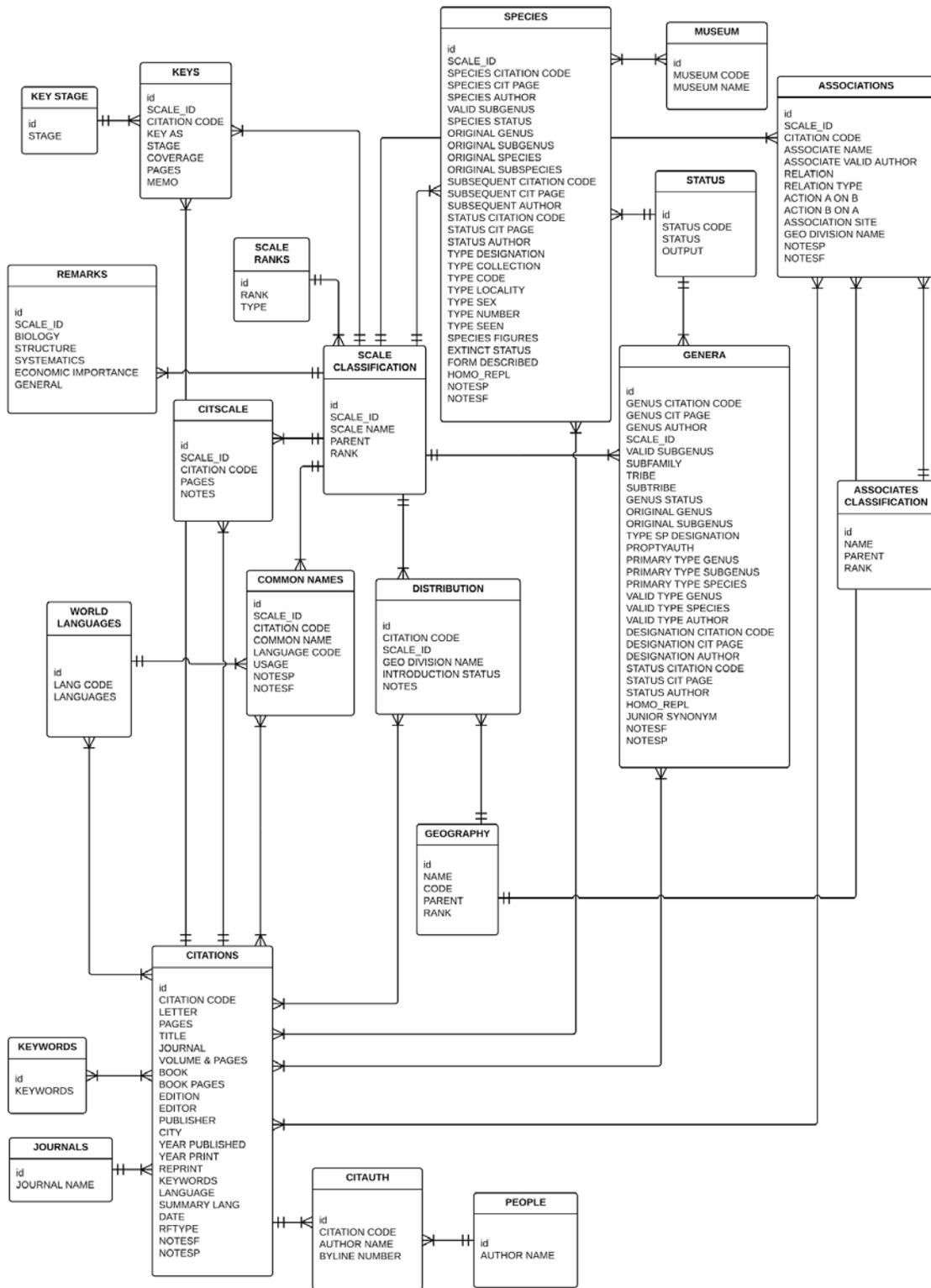


Figure 1. Diagram of the ScaleNet data model. The central tables to the database are the Citations, Scale Classification, Species and Genera tables. Together these tables validate the currently accepted valid names of scale insects, which are then used throughout the database to track ecological associations, distributions, taxonomic keys, etc. The figure depicts relationships between the tables using Crow's Foot Notation. The symbol || represents a one-and-only-one relationship. The crow's foot symbol represents a one-or-many relationship. Relationships can be asymmetrical, and the nature of the relationship of object A to object B is specified at the connection with B. For example, the relationship between Keys and Keys Stages would be read as 'One key can have one and only one key stage; a key stage can be in one or many keys.'

Results and Discussion

Database overview

Currently, ScaleNet contains 23,477 bibliographic records, pertaining to 9,509 currently valid scale insect names (8,194 of which are species combinations). Complete nomenclatural histories are available for each genus and species name, and ScaleNet associates 1,955 common names with 1,161 valid scientific names. Because of the agricultural importance of scale insects, the ScaleNet information about ecological associations and geographic ranges are particularly rich. There are 47,341 records of ecological associations between scale insects

and their hosts, natural enemies and mutualists. The geographic ranges of scale insect species are described by 32,641 records of occurrence in specific geopolitical or zoogeographic regions.

User interface

The public user interface exposes five major queries: (i) In the catalog query, users submit an available genus or species name to retrieve all of the information in the database associated with the valid form of that name. According to the rules of zoological nomenclature, a valid name is defined as the oldest available name for a genus or species, i.e. the one that has priority. An available species name is defined as any published binomial that is linked to a type specimen, and an available genus name is any published name that is linked to a type species. Users entering any of the available names associated with a species or genus will retrieve the data for the current valid name. The returned data view presents a nomenclatural history, lists of ecological associates and geopolitical units in which the taxon occurs, remarks on economic importance, biology, systematics and morphology and a complete bibliography. (ii) The places query allows users to retrieve a checklist of all of the scale insect species known to occur in a specified geopolitical or zoogeographic region. It is possible to constrain these searches to particular scale insect subgroups, e.g. specific genera. (iii) The ecological associates query returns a list of scale insect species associated with a specified host plant, natural enemy or mutualist. As in the places query, the results can be constrained to a scale insect subgroup. (iv) The references query gives users the ability to search the scale insect literature by author, year and keywords. (v) The common names query helps users make the connection between

common and scientific names of scale insect species. Users can also peruse the taxonomic diversity of scale insects and access catalogs by drilling down (and up) through a searchable scale insect classification.

Administrative interface

Previously, ScaleNet data were managed through FoxPro desktop clients. In the new version of ScaleNet, data management is through online administrative interfaces that Django automatically generates from the model metadata. The new ScaleNet affords curators considerably more flexibility in terms of where they work on ScaleNet. It is also more flexible in terms of who can manage the data. For most of its history, the ScaleNet curators were Y. Ben-Dov and D.R. Miller. Currently, ScaleNet is curated primarily by B.D. Denno. For a period of time following the retirement of Y. Ben-Dov and D.R. Miller, no one maintained the ScaleNet data. By the time B.D. Denno started her tenure as curator, ScaleNet was several years out of date, and many known data errors had gone uncorrected. Should there be a period in the future in which no one is able to assume a major responsibility for the curation of ScaleNet data, it may be possible to open the administrative interface up to the community of scale insect workers at large.

Data curation

Scale insect papers are added to ScaleNet after they have been identified through weekly Internet searches, or have been sent directly to ScaleNet curators by authors. Updates

to the database will be performed on a monthly basis. We aim for ScaleNet to include all published papers that deal with scale insects, but data entry is prioritized by subject, with the top priority going to papers that deal with the taxonomy and systematics of scale insects. Once a paper has been added to ScaleNet, curators extract information from that paper about the biological diversity of scale insects, and use that paper as a validation source for new records in various ScaleNet data classes (e.g. species names and geographic distributions). ScaleNet is meant to be a faithful representation of the literature; as a result, the data in ScaleNet is only as good as the data in the published literature. For the most part, ScaleNet curators do not judge the quality of the published information. If published information is erroneous, it needs to be corrected in a subsequent publication before that error will be corrected in ScaleNet.

Nevertheless, ScaleNet curators may exercise their discretion on issues of nomenclature and classification. Nomenclature changes in ScaleNet must comply with the International Code of Zoological Nomenclature, and if a taxonomic paper fails to do so, the proposed changes will not be committed to ScaleNet. Furthermore, ScaleNet is a comprehensive resource for a global fauna. It may be impossible for ScaleNet curators to commit published changes to scale insect systematics that apply to non-monophyletic groups, e.g. a family-level reclassification of only the Palearctic species of a global radiation.

The Future of ScaleNet

One impetus for the normalization of the ScaleNet data model was to increase the quality of the data through structural validations. However, because these validations were

lacking in the original application, a considerable amount of the data was invalid, and failed to be successfully migrated to the new platform. At the time of writing, manual restructuring and addition of these data is underway. Another impetus was to make ScaleNet more easily extendable, that is, increase the kinds of information accessible through ScaleNet. Some of what ScaleNet models, e.g. geographic ranges, can be more accurately modeled from specimen data, i.e. the metadata associated with physical insect specimens within natural history collections. Increasingly, these specimen data are available through web resources, such as the Global Biodiversity Information Facility's data portal (<http://www.gbif.org/>) and Discover Life (<http://www.discoverlife.org/>). In the past few years, data from hundreds of thousands of hemipteran specimens held in non-federal insect collections in the USA have been digitized by the Tri-Trophic Database project, an NSF-funded effort in the Advancing Digitization of Biological Collections program. In the future, we aim to include specimen-level data in ScaleNet's characterizations of scale insect biology. ScaleNet is used heavily by insect identifiers as a diagnostic tool. The extreme invasiveness of scale insect species stems in part from high propagule pressure, i.e. the sheer number of individuals which are brought along with plant materials to ports of entry. Scale insect species identifications are among the highest volume and most difficult jobs performed by inspection services. In the future we plan to make ScaleNet more useful as a diagnostic aid, by adding habitus images, taxonomic illustrations and complete taxonomic descriptions to catalog entries. ScaleNet is used increasingly by ecologists and evolutionary biologists. For example, recent studies have used ScaleNet data to address questions about the evolution of parthenogenesis (24) and diet breadth (25, 26). To facilitate the

compilation of comparative datasets from ScaleNet data, we plan to develop a ScaleNet web service API, i.e. more machine-friendly mechanisms for getting information from ScaleNet.

Chapter 3: Natural enemy pressure drives diversification in plant-feeding insects

Background

Insects are diverse. However, that diversity is not spread evenly across the insect phylogeny. Clades of plant-feeding species tend to be especially species rich; in fact, roughly half of all metazoan species are plant feeding insects (27). Why are there so many species of plant-feeding insects? In their seminal Escape and Radiate Hypothesis, Ehrlich and Raven (1964) proposed that the species diversification of plant-feeding insects and their hosts is driven by co-evolutionary adaptive radiation which hinges on the evolution of plant defensive chemistry (6). To date, the strict co-evolutionary dynamics of the Escape and Radiate Hypothesis are unsupported (for a review see 28), but following Ehrlich and Raven (1964), the theory of ecological speciation in plant-feeding insects has continued to fixate on host-plant interactions (29, 30).

Although ideas about the species diversification of plant-feeding insects have focused on host-plant interactions, there is considerable evidence that plant-feeding insect population dynamics are strongly affected by interactions with their natural enemies (11, 31, 32). There is also considerable evidence that plant-feeding insects take natural enemy pressure into account when making decisions about host use; that is, insects will often choose to utilize host plants that offer some form of protection or escape from their predators/parasitoids even if that comes at a cost of inferior food quality (12-15,33). Over evolutionary time, these interactions between

plant-feeding insects and their natural enemies could have a profound effect on species diversification.

Singer and Stireman (2005) formalized these considerations in their tri-trophic niche concept: host use in plant-feeding insects is the result of simultaneous optimization of nutrition and natural enemy exposure. Singer and Stireman (2005) also used the tri-trophic niche concept to make predictions about the dynamics of speciation in plant-feeding insects. They identified pressure from natural enemies as an important source of divergent selection on plant-feeding insect populations. Furthermore, they argued that host-use constraints would limit the ability of a plant-feeding insect lineage to respond to that divergent selection and escape natural enemy pressure. Therefore, they predicted that plant-feeding insect speciation rates would be positively correlated with 1) natural enemy pressure, and 2) ecological opportunity, i.e., the potential of escaping that pressure. For plant-feeding insects, escaping natural enemy pressure might be most easily achieved by switching host-plants. As follows, plant-feeding insect lineages with the highest speciation rates should be those under the greatest pressure from natural enemies, and with the weakest constraints on host use evolution.

Here we test these predictions in two diverse groups of plant-feeding insects, leafroller moths (Lepidoptera: Tortricidae) and scale insects (Hemiptera: Coccoidea). A major impediment to testing the macro-evolutionary predictions of the tri-trophic niche concept is that the natural enemy associations of most plant-feeding insect species are not well documented. Leafroller

moths and scale insects are exceptions. Each group includes many agricultural pests, and as a result their ecological associations have been studied extensively. Another factor that makes these groups particularly amenable for this study is that their species display considerable variation in host range, and considerable lability in their evolutionary associations with host plants. We use comparative phylogenetic methods to estimate the correlations between 1) the species richness of leafroller moth and scale insects clades, 2) the breadth of evolutionary constraints on host use in those clades, i.e., host range, and 3) the diversity of their parasitoid assemblages. We then estimate the effect of natural enemy pressure and host range, independently, on speciation rates.

Methods

Phylogeny

Tortricidae

We used PHLAWD (34) to assemble a multi-locus data set from published DNA sequences for phylogeny estimation. We assembled a supermatrix from alignments of sequences from 15 loci: (*28S*, *acetyl-CoA carboxylase*, *ala-tRNA*, *CAD*, *COI*, *DDC*, *glucose phosphate isomerase*, *glucose phosphate dehydrogenase*, *his-tRNA synthetase mRNA*, *nucleolar cysteine-rich protein*, *putative enolase protein*, *period mRNA*, *proteasome subunit mRNA*, *triosephosphate isomerase*, and *wingless*). This supermatrix had data for 815 species of leafroller moths. We used NCBI taxonomic classifications to annotate each species with their subfamily classification. We used RAxML (35) to estimate a maximum likelihood phylogeny,

under a GTR + CAT model of nucleotide substitution with parameters estimated independently for each locus. For our phylogeny searches, we constrained the estimate of relationships among 815 species to conform to the subfamily level relationships recovered by Regier et al 2012 (36). The maximum likelihood (ML) tree search had two phases. First, we estimated trees from 100 non-parametric bootstrap (BS) replicates of the supermatrix, using maximum parsimony starting trees. Second, we used each fifth BS tree as the starting tree for an ML search on the real supermatrix. We used treePL (37) to estimate divergence times, using a smoothing factor of 100, as selected by a cross-validation procedure. Optimization parameters were determined with treePL's prime command. For the divergence time estimates, we fixed the root age to 65 Ma (38) and applied a minimum age constraint of 33.9 Ma on the crown age of the subfamily Olethreutinae, based on fossil data (39). Taxa subtended by aberrantly long branches in the unrooted ML tree estimate were pruned from the dataset prior to divergence time estimation, as we interpreted these branches as indicative of potential problems with NCBI data curation, or phylogenetic supermatrix construction.

Coccoidea

For scale insects, we used a phylogeny published by Hardy et al. (2015), which was available in a Dryad repository (<http://dx.doi.org/10.5061/dryad.925cb>) (40). Briefly, this is a fossil-calibrated Bayesian estimate of the phylogenetic relationships and divergence times among 472 scale insect species based on a DNA sequence dataset of five loci (*COI*, *COII*, *EF1a*, *18S*, *28S*).

Trait Data

We used the taxonomic richness of parasitoid assemblages as a measure of specialized natural enemy pressure. Parasitoid assemblage species richness is positively related to parasitism rates (41, 42) so this measure should correlate with the total magnitude of specialized natural enemy pressure. Expressing specialized natural enemy pressure as the taxonomic richness of parasitoid assemblages should also correlate with the total diversity of divergent selective pressures. We used the taxonomic richness of host plants as a measure of ecological opportunity, that is, the potential of a plant-feeding insect lineage to escape natural enemy pressure via host switching. We modeled each trait -- natural enemy pressure and host range -- as counts of taxa at two levels of classification: family and genus.

We collected natural enemy records for 1,206 species of leafroller moths. These were aggregated from two databases: “Universal Chalcidoidea Database” (43), maintained by the Natural History Museum in London, and the online catalogue “Arthropods of Economic Importance: Eurasian Tortricidae” (44) hosted by the Netherlands Biodiversity Information Facility. We also collected host range records for 2,374 leafroller moth species from the “Food Plant Database for the Leafrollers of the World (Lepidoptera: Tortricidae)” (45). In total, we were able to find 1,603 unique moth-natural enemy interactions, and 10,530 unique moth-host plant interactions. For scale insects, we downloaded ecological interaction data from ScaleNet (5). We collected natural enemy data for 716 scale insect species and host plant records for 7,303

scale insect species, representing 4,654 unique scale-natural enemy interactions and 46,564 unique scale-host plant interactions. We standardized the names of moth species according to the “Online World Catalogue of the Tortricidae” (46). The names of all ecological associates (natural enemies and host plants) were standardized with the Taxonomic Name Resolution Service (47) with 70% fuzzy matching.

Comparative analysis

We used the macro-caic function in the R (48) package ‘caper’ (49) to fit linear models in which species richness was the response variable, and our parameterizations of natural enemy pressure and host range were predictors. We used Akaike’s Information Criterion (AIC) scores to compare the fit of a tri-trophic model to the fit of host-only and natural-enemy-only biotrophic models. Next, we used ML optimization of Quantitative State Dependent Speciation and Extinction (QuaSSE) (50) models to estimate the effect of natural enemy pressure and host range on speciation rates. We fit three models: 1) a model in which speciation and extinction rates are constant and independent of ecological traits (natural enemy pressure or host range), 2) a model in which speciation rates are a linear function of natural enemy pressure, and 3) a model in which speciation rates are a linear function of host-plant range.

The final dimensions of each comparative dataset varied among analyses. For scale insects, the macro-caic comparative dataset used a phylogeny relating 184 genera, 112 of which had natural enemy data and 184 of which had host data (Figure 2a). For leafroller moths,

the macro-caic analyses were performed on a phylogeny relating 185 genera, 74 of which had NE data and 184 of which had host data (Figure 2b). For the QuaSSE analyses, we pruned species level phylogenies to include only tips with ecological association data. We had natural enemy data for 121 leafroller moth species and host use data for 473 species. For scale insects, we had natural enemy data for 192 species and host range data for 417 species.

Recent work has shown that SSE models have a high Type I error rates when used with empirical phylogenies, possibly due to complexities of real phylogenetic processes that are poorly captured by standard Equal Rates Markov models (e.g., the birth-death model) (51). To calibrate critical values for significance testing, we used simulation of diversification-neutral traits to create a null distribution of the differences in AIC scores between models. Thus the significance of each empirical model comparison was assessed by comparing the AIC difference to the distribution of AIC differences (dAIC) calculated for models estimated from 100 neutral simulations.

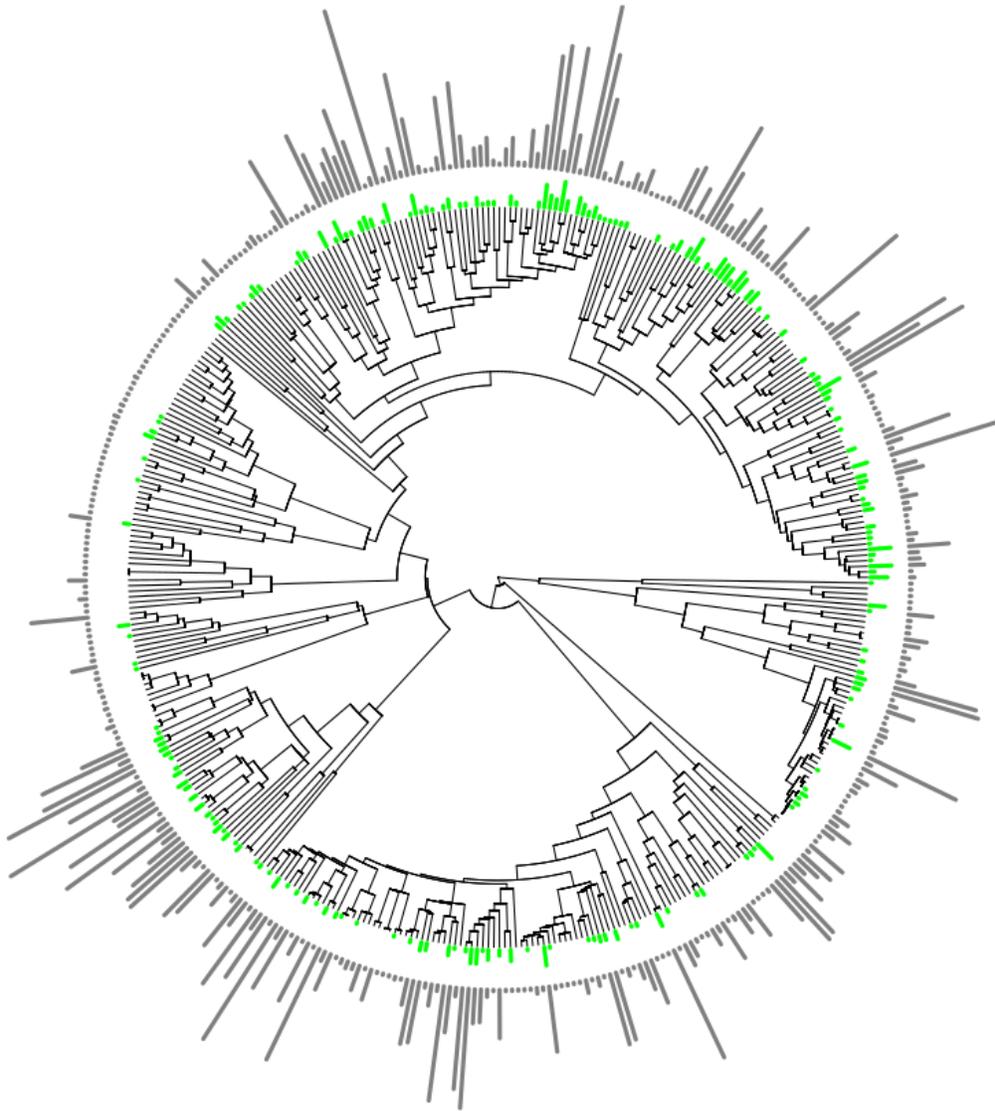


Figure 2. Phylogeny of 472 scale insect species by Hardy et. al 2015. Each species has been annotated with information for its ecological associates. The length of each bar represents counts of family-level taxa (green = natural enemies; grey = host plants). NE ranges from 1-20. HR ranges from 1-129.

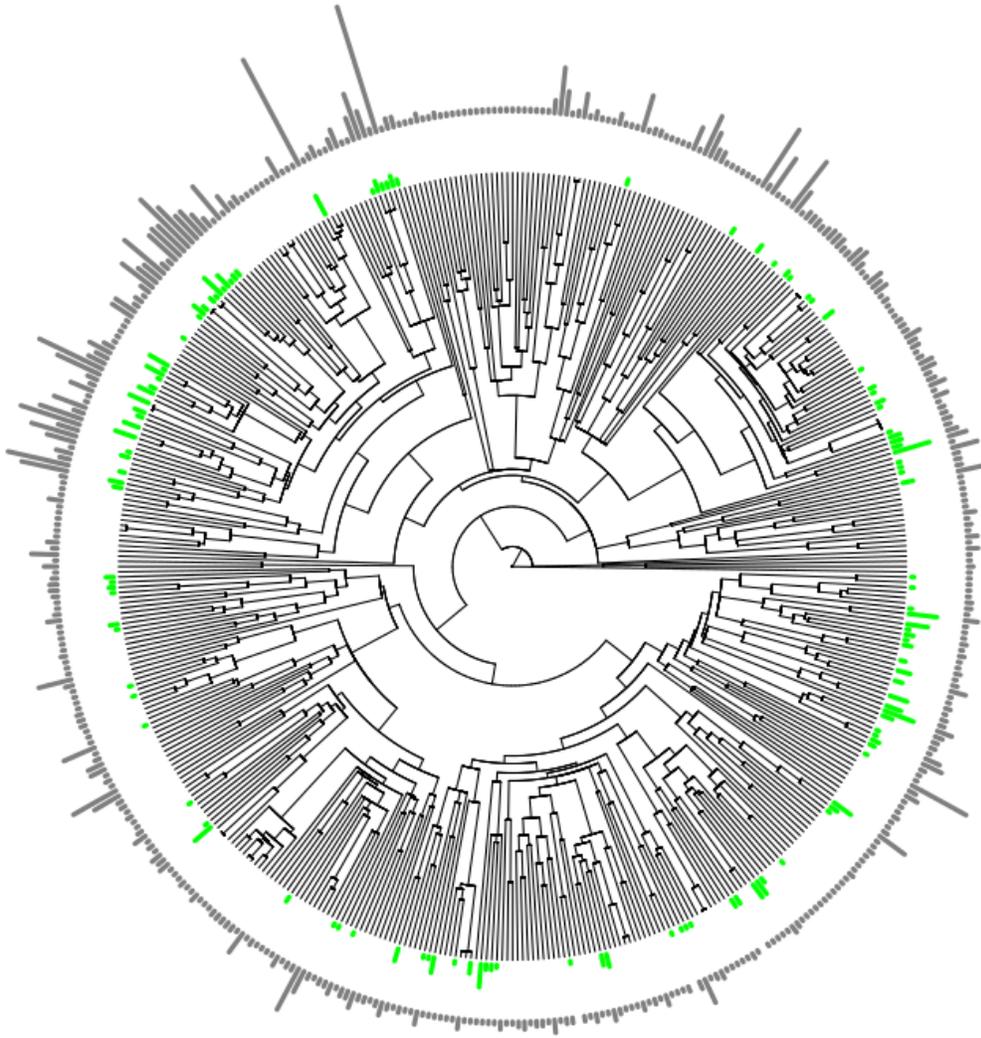


Figure 3. Phylogeny of 852 leafroller moth species. Each species has been annotated with information for its ecological associates. The length of each bar represents counts of family-level taxa (green = natural enemies; grey = host plants). NE ranges from 1-14. HR ranges from 1-51.

Results

For brevity, we mainly discuss the results of analyses in which host ranges and natural enemy pressures were parameterized as counts of family-level taxa. In almost all cases, the

results of analysis based on counts of genera were qualitatively the same; that is, our results were robust to the taxonomic level that was counted as measure of ecological breadth. Where there was a difference between parameterizations, we mention it below.

Effects on species richness

The tri-trophic model was the best fit to the scale insect data; both natural enemy pressure and host range had significant positive effects on scale insect species richness (Table 1, Figure 4a). The effects of host range and natural enemy pressure on scale insect species richness were also positive and significant in each of the bi-trophic models, although the magnitude of the natural enemy effect was greater than the host range effect (Table 1). The tri-trophic model was favored over the bi-trophic natural enemy model by ~2.5 AIC units.

For the leafroller moth dataset, the bi-trophic natural enemy model was the best fit. However, the AIC score of the tri-trophic model and the bi-trophic natural enemy model were essentially indistinguishable (Table 1, Fig 4b). In the best-fitting model, natural enemy pressure had a positive and significant effect on leafroller moth species richness. In the tri-trophic model, there was a positive and significant effect of natural enemy pressure on species richness and a positive but nonsignificant effect of host range. Both bi-trophic models had positive and significant effects on species richness (Table 1).

Model	Effect	Std.Error	AIC
Scale Insects - Family level			
Tritrophic (NE HR)	0.33* 0.033*	0.14 0.015	487.27
Bitrophic - NE	0.43**	0.013	490.03
Bitrophic - HR	0.069***	0.13	818.56
Scale Insects - Genus level			
Tritrophic (NE HR)	0.11* 0.012*	0.055 0.0058	484.68
Bitrophic - NE	0.18***	0.047	812.18
Bitrophic - HR	0.026***	0.0043	486.80
Leafroller moths- Family level			
Tritrophic (NE HR)	0.97** 0.104	0.35 0.08	279.77
Bitrophic - NE	1.21***	0.30	279.49
Bitrophic - HR	0.36***	0.066	774.75
Leafroller moths - Genus level			
Tritrophic (NE HR)	0.24** 0.043	0.088 0.034	279.60
Bitrophic - NE	0.30***	0.073	279.28
Bitrophic - HR	0.15***	0.03	776.55

Table 1. Results from tri-trophic and bi-trophic caper models for each herbivore group. The response variable for all models is herbivore genera diversity. The explanatory variables are natural enemies and host range (tri-trophic), natural enemies (bitrophic), or host range (bitrophic). Each trait is a count of either genus or family level taxa. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

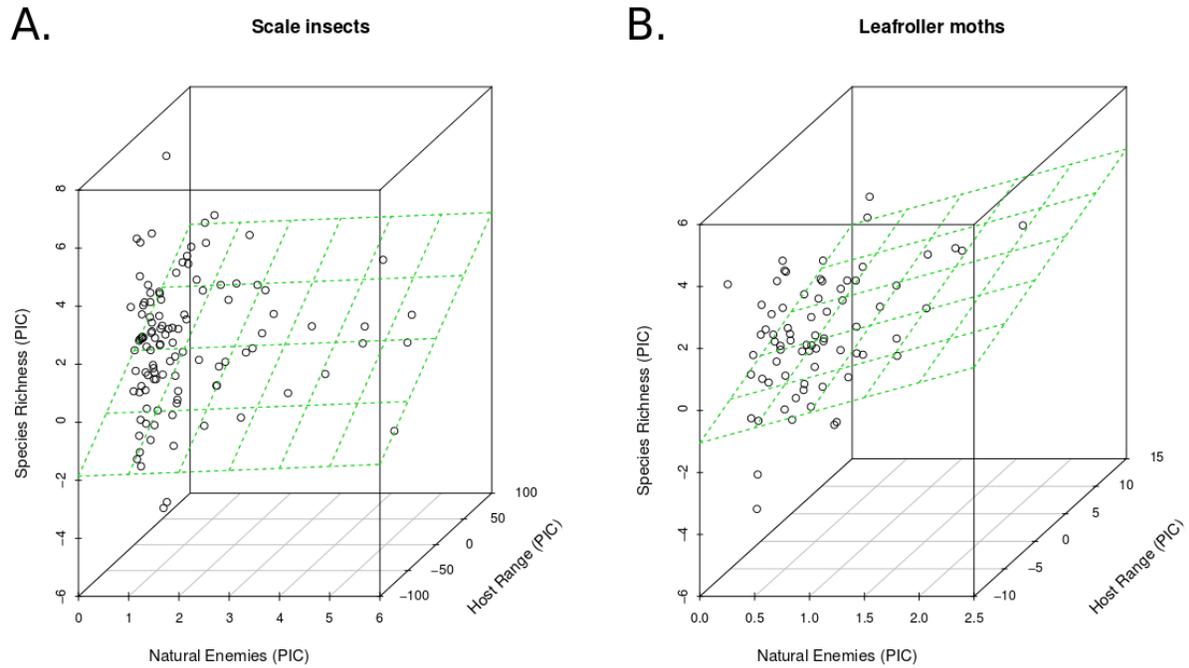


Figure 4. Scatterplot for a tritrophic model for a) scale insects and b) leafroller moths. The green grid represents the regression plane for the model. Measures of each axis is phylogenetic independent contrasts (PIC) calculated by caper.

Effects on speciation rates

For leafrollers, we recovered significant positive effects of natural enemy pressure ($p < 0.1$) and host range ($p < 0.1$) on speciation rates (Table 2). For scale insects, we also recovered a positive effect of natural enemy pressure ($p < 0.1$) and host range ($p < 0.1$) when measured

as counts of genera. The effect of host range was not significant when the trait was modelled as counts of families (Table 2).

Model	Intercept	Effect	dAIC
Scale Insect- Family level			
Natural Enemies	0.65	-0.0034	2.31
Host Range	0.043	0.00049	14.57???
Scale Insects- Genus level			
Natural Enemies	0.502	0.0019	52.81***
Host Range	0.36	0.00071	76.48
Leafroller moths- Family level			
Natural Enemies	0.17	0.014	1044.62***
Host Range	0.25	0.0048	92.30***
Leafroller moths- Genus level			
Natural Enemies	0.15	0.0065	1051.46***
Host Range	0.53	0.0025	23.57**

Table 2. Results from bitrophic QuaSSE models. In each model, speciation rates are a linear function of either natural enemies or host range. Scores in the column dAIC are the difference in AIC scores between a constant model and a linear model. The statistical significance of these differences is measured from a null distribution of AIC scores from simulated neutral-trait evolution models. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Discussion

In this study, we have recovered the first empirical support for the macroevolutionary predictions of the tri-trophic niche concept. Specifically, Singer and Stireman (2005) predicted that speciation rates in plant-feeding insects will be positively related to the amount of divergent selective pressure applied by natural enemies, and the potential to escape that pressure by evolving host use. The patterns of species richness, host range, and natural enemy pressure in the phylogenies of scale insects and leafroller moths are consistent with these predictions. Moreover, our comparative phylogenetic analyses suggest that natural enemy pressure may actually have a stronger influence on plant-feeding insect diversity than host-plant range *per se* -- the estimated model coefficients were an order of magnitude greater for natural enemy pressure than for host ranges. This is a surprising result, and suggests that there is a limit to what we can learn about ecological divergence and speciation of plant-feeding insect lineages from models that only account for host-plant interactions.

Given this result, how do we explain the previous success of host-plant bi-trophic models in predicting patterns of species richness in plant-feeding insect phylogenies? For example, comparative phylogenetic analyses support the Musical Chairs Hypothesis prediction that rates of speciation should be correlated with rates of host switching (9). The explanation likely lies with the fact that so much of a plant-feeding insect species' ecology is tied to host plants. In addition to being a food source, host plants are targets for oviposition and mate finding. They

are also the place in which natural enemy interactions occur, and different plant species may confer variable levels of resistance to or refuge from natural enemy pressures. The Musical Chairs Hypothesis does not explicitly state why plant-feeding insect lineages should switch hosts, but rather assumes that selective environments are dynamic and niche optimization will often entail host switching. Natural enemy interactions could be a crucial factor in determining those selective environments and spurring host switching as well as speciation (52).

Singer and Stireman (2005) explicitly focused on specialized natural enemies, especially parasitoids, as agents of divergent selection on plant-feeding insects that could drive speciation. In their view, generalist natural enemies are ubiquitous, and apply pressure to plant-feeding insects in ways that would not be abated by the evolution of host-plant use. In the present study, we found that exclusively specialized natural enemy pressure had a strong positive effect on the species diversity of leafroller moths. By contrast, for scale insects our measures of natural enemy pressure encompassed both specialist and generalist natural enemies, i.e., parasitoids and a variety of predators. The strong positive effect of natural enemy pressure on scale insect species diversity could be interpreted in two ways. One, it could indicate that the signal for specialist natural enemies was strong enough to dominate the effects. Two, it could indicate that pressure from generalist natural enemies also contribute to ecological divergence and speciation in plant-feeding insects, contrary to the expectations of Singer and Stireman (2005).

While the ecological data available for our focal plant-feeding insect taxa are exceptionally rich, we were only able to build a suitable comparative dataset covering about 1% of extant leafroller moth species and 8% of extant scale insect species. Host plant information is much easier to come by and, in the near term, will likely remain the focus of explorations of ecological macroevolution in plant-feeding insects. Nevertheless, this study underscores the importance of developing more complex models of the niches of plant-feeding insects, and in particular of more thoroughly documenting the interactions plant-feeding insects have with their natural enemies.

Conclusions

The phylogenetic patterns of the evolution of host range and natural enemy pressure are consistent with macroevolutionary predictions of the tri-trophic niche model for plant-feeding insects. Natural enemy pressure appears to have an even greater effect on the species richness and speciation rates of plant-feeding insect lineages than the breadth of their host-plant associations. These results emphasize the importance of using more complex models of the niches of plant-feeding insects in studies of their ecology and ecological evolution.

Chapter 4: Conclusions

What is the true scope of the taxonomic and ecological diversity of plant-feeding insects? In this thesis, I describe my work using open-source software tools to rebuild the database ScaleNet, one of the richest and most comprehensive characterizations of the biodiversity of a plant-feeding insect group. This project has created a platform that will foster the efficient growth of our knowledgebase for scale insect species diversity. As a result, an understanding of the true scope of the taxonomic and ecological diversity of scale insects is much closer at hand than it is for almost any other plant-feeding insect group of comparable size. This positions scale insects to become an important model system for studies of insect evolutionary ecology (e.g., 53).

What factors drive plant-feeding insect speciation? In this thesis, I combined information about the ecological associations of scale insects and leafroller moths (from ScaleNet and a variety of other sources) with estimates of the phylogenetic histories of those groups in comparative phylogenetic analyses. I found that natural enemy pressure is an important predictor of species richness, and speciation rates, in scale insects and leafroller moths. This represents the first empirical support for the macroevolutionary predictions of the tri-trophic niche model for plant feeding insects. Natural enemy interactions are an important component of the ecology of plant-feeding insects, and their future inclusion into models that attempt to explain plant-feeding insect diversity will be necessary.

The skills I developed through this research have put me in a better position to continue exploring my academic interests in the diversity of insects. In my future research plans, I am particularly interested in how natural enemy interactions affect spatial and temporal variation in plant-feeding insect populations. Additionally, it will be interesting to analyze the effect of natural enemy pressure on the evolution of host use in plant-feeding insects. Lastly, it will be important to consider the roles of the host specificity of natural enemies in conditioning the effects that they have on plant-feeding insect ecology, and evolution.

References

1. Pimm SL, et al. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344.6187: 1246752.
2. Footitt RG, Adler PH. (2009) *Insect biodiversity: science and society*. John Wiley & Sons.
3. Erwin, Terry L. (1982) Tropical forests: their richness in Coleoptera and other arthropod species." *Coleopterists Bulletin* 36.1: 74-75.
4. Mora, Camilo, et al. (2011) How many species are there on Earth and in the ocean?. *PLoS Biol* 9.8: e1001127.
5. García Morales M, et al. (2016) ScaleNet: A literature-based model of scale insect biology and systematics. Database. doi: 10.1093/database/bav118. <http://scalenet.info>.
6. Ehrlich PR, Raven PH. (1964) Butterflies and Plants: A Study in Coevolution. *Evolution*, 18(4): 586-608.
7. Futuyma DJ, Agrawal AA. (2009) Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences*, 106.43 : 18054-18061.
8. Janz N, Nylin S. (2008) The oscillation hypothesis of host-plant range and speciation. *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects*, 203-215.
9. Hardy NB, Otto, SP. (2014) Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society of London B: Biological Sciences*, 281.1795: 20132960.
10. Symondson WOC, Sunderland KD, Greenstone MH. (2002) Can generalist predators be effective biocontrol agents? *Annual review of entomology*, 47.1: 561-594.
11. Singer MS, Stireman, JO. (2005) The tri-trophic niche concept and adaptive radiation of phytophagous insects. *Ecology Letters*, 8(12): 1247-1255.
12. Singer MS, et al. (2004) Disentangling food quality from resistance against parasitoids: diet choice by a generalist caterpillar. *The American Naturalist*, 164.3: 423-429.
13. Mira A, Bernays EA. (2012) Trade-offs in host use by *Manduca sexta*: plant characters vs natural enemies. *Oikos*, 97(3): 387-397.
14. Diamond SE, Kingsolver JG. (2010) Fitness consequences of host plant choice: a field experiment. *Oikos*, 119(3): 542-550.
15. Nomikou M, Janssen A, Sabelis MW. (2003) Herbivore host plant selection: whitefly learns to avoid host plants that harbour predators of her offspring. *Oecologia*, 136(3): 484-488.
16. Douglas AE. (2006) Phloem-sap feeding by animals: problems and solutions. *J. Exp. Bot.*, 57 (4), 747-754.

17. Styrsky JD, Eubanks MD. (2007) Ecological consequences of interactions between ants and honeydew-producing insects. *Proc. R. Soc. B*, 274, 151-164; doi: 10.1098/rspb.2006.3701.
18. Stadler B, Michalzik B, Muller T. (1998) Linking aphid ecology with nutrient fluxes in a coniferous forest. *Ecology*, 79(5), 1514-1525.
19. Miller DR. (2005) Introduced scale insects (Hemiptera: Coccoidea) of the United States and their impact on U.S. agriculture. *Proc. Entomol. Soc. Wash.*, 107(1), 123-158.
20. Normark, BB (2003) The evolution of alternative genetic systems in insects. *Annu. Rev. Entomol.*, 48, 397-423.
21. Miller DR, Ben-Dov Y, Gibson GAP. (1999) ScaleNet: A Searchable Information System on Scale Insects, *Entomologica*, 33, 37-46.
22. Ben-Dov Y. (1994) A systematic catalogue of the mealybugs of the world (Insecta: Homoptera: Coccoidea: Pseudococcidae: Putoidae). Intercept Ltd., Andover, UK, 686 pp.
23. Ben-Dov Y. (1993) A systematic catalogue of the soft scales of the world (Insecta: Homoptera: Coccoidea: Coccidae). *Flora and Fauna Handbook No. 9*, Sandhill Crane Press, Gainesville, Florida, USA, 536 pp.
24. Ross L, et al. (2012) The role of endosymbionts in the evolution of haploid-male genetic systems in scale insects (Coccoidea). *Ecology and Evolution*, 2(5), 1071-1081.
25. Lin YP, et al. (2015) Does host-plant diversity explain species richness in insects? A test using Coccidae (Hemiptera). *Ecological Entomology*, 40(3), 299-306. doi: 10.1111/een.12191
26. Peterson DA, et al. (2015) Phylogenetic analysis reveals positive correlations between adaptations to diverse hosts in a group of pathogen-like herbivores. *Evolution*, 69(n/a). doi: 10.1111/evo.12772
27. Mitter C, Farrell B, Weigmann, B. (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *American Naturalist*, 132:107 – 128.
28. Futuyma DJ, Agrawal AA. (2009) Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences*, 106(43): 18054-18061.
29. Nylin S, Slove J, Janz N. (2014) Host plant utilization, host range oscillations and diversification in nymphalid butterflies: a phylogenetic investigation. *Evolution*, (1): 105-124.
30. Janz N, Nylin S, Wahlberg N. (2006) Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology*, 6.1: 1.
31. Hawkins BA, Cornell HV, Hochberg ME. (1997) Predators, parasitoids, and pathogens as mortality agents in phytophagous insect populations. *Ecology*, 78.7: 2145-2152.
32. Brown JM, et al. (1995) The role of natural-enemy escape in a gallmaker host-plant shift. *Oecologia*, 104(1): 52-60.

33. Bernays E, Graham M. (1988) On the evolution of host specificity in phytophagous arthropods. *Ecology*, 69.4: 886-892.
34. Smith SA, Beaulieu JM, Donoghue, MJ. (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC evolutionary biology*, 9.1: 37.
35. Stamatakis, A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies". *Bioinformatics*.
36. Regier JC, et al. (2012) A molecular phylogeny for the leaf-roller moths (Lepidoptera: Tortricidae) and its implications for classification and life history evolution. *PLoS One*, 7(4): e35574.
37. Smith SA., O'Meara BC. (2012) treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics*, 28(20): 2689-2690.
38. Misof B, et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210): 763-767.
39. Sohn JC, et al. (2012) An annotated catalog of fossil and subfossil Lepidoptera (Insecta: Holometabola) of the world. *Zootaxa*, 3286: 1-132.
40. Hardy NB, Peterson DA, Normark, BB. (2015) Scale insect host ranges are broader in the tropics. *Biology letters*, 11(12): 20150924.
41. Macfadyen S, et al. (2011) Parasitoid diversity reduces the variability in pest control services across time on farms. *Proceedings of the Royal Society of London B: Biological Sciences*, 283(1826): .
42. Peralta G, et al. (2014) Complementarity and redundancy of interactions enhance attack rates and spatial stability in host–parasitoid food webs. *Ecology*, 95(7): 1888-1896.
43. Noyes, J.S. (2015) Universal Chalcidoidea Database. World Wide Web electronic publication. <http://www.nhm.ac.uk/chalcidoids>
44. Meijerman L, Ulenberg SA. (2004) Eurasian Tortricidae. *Arthropods of Economic Importance*.
45. Brown JW, Robinson G, Powell JA. (2008) Food plant database of the leafrollers of the world (Lepidoptera: Tortricidae) (Version 1.0). <http://www.tortricid.net/foodplants.asp>.
46. Gilligan T M, Baixeras J, Brown JW, Tuck KR. (2014) T@RTS: Online World Catalogue of the Tortricidae (Ver. 3.0). <http://www.tortricid.net/catalogue.asp>.
47. Boyle B, et al. (2013) The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics*, 14:16. doi:10.1186/1471-2105-14-16
48. R Core Team. (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL <https://www.R-project.org/>.

49. Orme, D. (2013) The caper package: comparative analysis of phylogenetics and evolution in R. R package version 5.2.
50. FitzJohn, RG. (2010) Quantitative traits and diversification. *Systematic biology*: syq053.
51. Rabosky DL, Goldberg EE. (2015) Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology*, 64(2): 340-355.
52. Drès M, Mallet J. (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 357(1420): 471-492.
53. Hardy NB, Peterson DA, Normark BB. (2015) Scale insect host ranges are broader in the tropics. *Biology letters*, 11.12: 20150924.