MACRO AND MICRO EVOLUTIONARY DETERMINANTS OF DIVERSIFICATION IN CECIDOMYIID FLIES

by

Jeffrey B. Joy
B.Sc. University of Victoria, 1996

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the
Department of Biological Sciences

© Jeffrey B. Joy 2009

SIMON FRASER UNIVERSITY

Fall 2009

All rights reserved. However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for *Fair Dealing*. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name:	**Jeffrey Bruce Joy**
Degree:	**Doctor of Philosophy**
Title of Thesis:	**Macro and micro evolutionary determinants of diversification in cecidomyiid flies**

Examining Committee:

Chair: **Dr. Harold Hutter**
Professor

____

Dr. Bernard Crespi
Senior Supervisor
Professor, Department of Biological Sciences, SFU

____

Dr. Arne Mooers
Supervisor
Associate Professor, Department of Biological Sciences, SFU

____

Dr. Felix Breden
Supervisor
Professor, Department of Biological Sciences, SFU

____

Dr. Michael Hart
Internal Examiner
Associate Professor, Department of Biological Sciences, SFU

____

Dr. Brian Farrell
External Examiner
Professor, Department of Organixmal and Evolutionary Biology
Harvard University

Date Defended/Approved: November 24, 2009

SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Last revision: Spring 09

Abstract

Understanding how diversification proceeds during adaptive radiation requires studies of diversity at multiple levels (within species, between species, and above the species level). Adaptive radiation involves both the radiation of species from a common ancestor and partitioning of environments by those species through ecological divergence. Phytophagous insects comprise the bulk of the world's biological diversity, and understanding the evolutionary processes that drive their diversification is a central theme in evolutionary biology. The ecologically specialized relationship between gall-inducing phytophagous insects and their host plants makes them ideal systems for examining causal mechanisms of evolutionary diversification. Gall midges (Diptera: Cecidomyiidae) are especially useful among gall-inducing insects because they are diverse, disperse over large distances, and have radiated among a variety of host-plant species; taxonomic data also show that many genera exhibit large groups of closely-related species on single host-plant species. This thesis examines the macro- and micro-evolutionary diversification of host-associated species of gall-inducing midges. Increases in cecidomyiid diversity between host-plant taxa were associated with increases in ecological opportunity, plant lineage age, and plant architectural complexity and decreases in plant insularity. Diversification of *Rhopalomyia* gall midges within plant family (Asteraceae) results from a combination of host-plant shifts and within host-plant speciation. Diversification of *Asphondylia* gall midges within a single host-plant species results from within host-plant speciation. Speciation without a host shift in both

Asphondylia and *Rhopalomyia* is associated with shifts among plant parts and shifts among time periods indicating that such shifts may be general mechanisms facilitating divergence within a single host-plant species. Divergence population genetics supports inferences of large ancestral population size and gene flow during divergence between a species pair shifted in life-history timing; and small ancestral population size and no gene flow during divergence in a species pair displaying divergence in plant-part use. Comparative analyses of *Asphondylia* wing length and ovipositor length suggest strong divergent selection on ovipositor length accompanies evolutionary shifts between host-plant parts. Studies of other radiations of cecidomyiids combined with analyses of genes putatively involved in the evolution of reproductive isolation will provide a more complete understanding of the evolutionary processes involved in cecidomyiid diversification.

Keywords: Adaptive radiation; speciation; phytophagous insects; evolutionary processes; transfer RNAs; Cecidomyiidae

Dedication

This thesis is dedicated to William Donald Hamilton for discussions on a myriad of topics, evolutionary and otherwise, in the depths of the rainforests of the Democratic Republic of Congo which inspired me to travel this path.

// Acknowledgements

"We are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size."

Bernard of Chartres – 12 Century

The work that comprises this thesis is deeply in the debt of my intellectual forebears and a product of the rich intellectual environment and encouragement provided by my advisor, committee, colleagues, friends, and family.

First and foremost it is difficult to convey the immensity of my gratitude towards my supervisor Bernie Crespi for his patience, support, sound advice, and constant source of scientific inspiration. Bernie's considerable skill as a supervisor is reflected (as one example) in his remarkable capacity to return manuscripts with insightful, thoughtful, constructive comments in remarkably short time frames measurable in hours rather than days or weeks.

Arne Mooers was a tireless supporter throughout my time at SFU, he was a valuable, effective and supportive member of my advisory committee providing useful feedback and actively participated in many of my intellectual endeavours both thesis and non-thesis related. Felix Breden was also a tremendously valuable member of my advisory committee encouraging the development of aspects of my work particularly as regards gene duplication and the evolution of diversity in gene families, provided useful

comments on manuscripts, and was generally supportive of the breadth of projects I undertook. Nancy Moran was an invaluable source of intellectual support and encouragement both during my time spent in Tucson and remotely. Nancy not only allowed me to fill a not insignificant portion of her laboratory with small vials of obscure flies but she also provided me with equipment in the field and generally supported and made my experience in Tucson as wonderful and smooth as possible. Mike Hart provided regular and invaluable feedback both on the various chapters of my thesis (especially as regards coalescent analyses) but also contributed to the broad range of other projects I undertook. Andy Beckenbach though not officially on my advisory committee was always there to lend advice along with the latest primers. Brian Farrell deserves mention not just for his role as my external examiner but also for the inspiration I drew from his work before we met.

Many people helped me in the field and in the lab; of these people Hana Kucera deserves special mention for enduring weeks in the desert with me and for invaluable assistance with work in the laboratory. Kristy Williams also helped immensely in the laboratory.

The value of the feedback from everyone in the research groups at SFU (FAB* lab) and University of Arizona (Moran-Ochman group) simply can't be over stated. The list of people who deserve mention for intellectual contributions is immense, in particular I thank Gerhard Gries, Regine Gries, Christine Parent, Stevan Springer, Patrik Nosil, Rutger Vos, Erica Jeffery, Graham Thompson, Michael Loeb, Howard Rundle, Jeremy Bono, Mick Elliot, Tanja Schwander, Devin Arbuthnott, Corey Watson, Laura Weir, Tommi Nyman, Boris Igic, Patrick Abbot, Evan Braswell, Betsy Arnold, Michael

Worobey, Tom Gilbert, Nicole Gerardo, Alex C.C. Wilson, Liliana Davalos, Gwen Waring, Ray Gagne, Art Borkent, and Netta Dorchin.

Many thanks are due to my parents and family for patience, and understanding. Timothy, Forrest, and Ruth Joy in addition to being supportive also collected material from both the Sonoran and Great Basin deserts. I thank also my many friends both science and non-science related and apologize to those who through oversight I have neglected to mention.

The SFU Department of Biological Sciences and the Interdisciplinary Research in the Mathematical and Computational Sciences (IRMACS) Centre provided world-class facilities for me to conduct my research while at SFU. The Natural Sciences and Engineering Council of Canada and the Society for Systematic Biology provided funding and I am thankful to the Smithsonian Institution for providing museum specimens of cecidomyiids.

Table of Contents

Approval ii
Abstract iii
Dedication v
Acknowledgements vi
Table of Contents ix
List of Figures xii
List of Tables xiv

Chapter 1: General introduction to adaptive radiation and diversification 1

1.1 Adaptive Radiation Among Host-plant Species 2
1.2 Adaptive Radiation Within Host-plant Species 2
1.3 Summary of Thesis Chapters 3
1.4 Literature Cited 6

Chapter 2: Island phytophagy 9

2.1 Abstract 9
2.2 Introduction 10
2.3 Methods 12
2.3.1 Host-plant preferences database 12
2.3.2 Statistical methods 12
2.3.3 Estimation of plant species richness diversity relationships 13
2.3.4 Inference of plant insularity and insect diversity 13
2.3.5 Plant taxon age 13
2.3.6 Plant complexity and insect species richness 14
2.4 Results 14
2.5 Discussion 16
2.5.1 Insularity 16
2.5.2 Ecological Opportunity (Plant Diversity) 17
2.5.3 Plant Lineage Age 18
2.5.4 Plant Architectural Complexity 19
2.5.5 Plant Abundance and Distribution 20
2.5.6 Conclusion 21
2.6 Acknowledgments 22
2.7 Literature Cited 22

Chapter 3: Radiation of *Rhopalomyia* gall midges within the plant family Asteraceae 38

3.1 Abstract 38
3.2 Introduction 39
3.3 Methods 43

3.3.1 Collection Sites 43
3.3.2 Collection of DNA Data 43
3.3.3 Phylogenetic Analyses...44
3.4 Results 45
3.4.1 Dataset 45
3.4.2 Phylogentic Inferences and Evolution of Host-plant Use 45
3.4.4 Ecological Contexts of Within-host Speciation 47
3.5 Discussion 48
3.6 Acknowledgements 51
3.7 Literature Cited 51

Chapter 4: Adaptive radiation of gall-inducing insects within a single host plant species..61

4.1 Abstract 61
4.2 Introduction..62
4.3 Methods..66
4.3.1 Collection Sites and Methods..66
4.3.2 Collection of DNA data 66
4.3.3 Phylogenetic Analyses..67
4.3.4 Comparative Analyses 69
4.4 Results 70
4.4.1 Dataset 70
4.4.2 Phylogenies...70
4.4.3 Evolution of Host Plant Use...71
4.4.4 Host-Plant Colonization Sequence...72
4.4.5 Ecological Adaptation to Specific Plant Parts...72
4.5 Discussion 73
4.5.1 Potential Mechanisms of Within Host-plant Speciation...74
4.5.2 Ecological Adaptation to Specific Plant Parts...78
4.5.3 Conclusions 79
4.6 Acknowledgments 80
4.7 Literature Cited 80

Chapter 5: Historical gene flow and population demographic history during adaptive radiation of gall-inducing insects 96

5.1 Abstract 96
5.2 Introduction..97
5.3 Methods..100
5.3.1 Collection Sites and Methods..100
5.3.2 Collection of DNA Data 101
5.3.3 Phylogenetic Analyses...101
5.3.4 Geographic Distribution of Ancient Alleles..102
5.3.5 Coalescent Gene Flow and Demographic Analyses 102
5.4 Results 104
5.4.1 Dataset 104
5.4.2 Phylogenetics..106
5.4.3 Geographic Distribution of Ancient Alleles..106
5.4.4 Coalescent Analyses 107

5.5 Discussion 108
5.5.1 Hypotheses of Divergence..110
5.6 Acknowledgments 113
5.7 Literature Cited 113

Chapter 6: Major Findings and Conclusions 124

6.1 Summary 124
6.2 Nexus of Macro and Micro Evolutionary Processes 125
6.3 Within Host Speciation: Plant-part Shifts 126
6.4 Within Host Speciation: The Importance of Time 127
6.5 Within Host Speciation - Generality?...129
6.7 Final Remarks 130
6.8 Literature Cited 130

Appendix 135

Appendix 1: Comparative mitochondrial genomics reveals truncated transfer RNA genes are a common feature of Cecidomyiid mitochondrial genomes 135
S1.1 Abstract 135
S1.2 Introduction..136
S1.3 Methods..139
S1.3.1 Source and Collection of Specimens..139
S1.3.2 Genome Sequencing 140
S1.3.3 Annotation of the Sequences...140
S1.3.4 Phylogeny Reconstruction and Character Evolution..........................141
S1.3.5 Data Deposition 142
S1.4 Results...142
S1.4.1 tRNA Gene Rearrangements..143
S1.4.2 Features of Cecidomyiid Mitochondrial Genomes..............................144
S1.4.3 Features of Cecidomyiid Mitochondrial Genomes..............................146
S1.4.4 Structural Characteristics of Cecidomyiid tRNAs...............................147
S1.4.5 Phylogenetic Results...148
S1.5 Discussion...149
S1.5.1 Rearrangement of tRNA Genes 150
S1.5.2 Truncated tRNA Genes 151
S1.5.3 Parallel Evolution of Truncated tRNA Genes 154
S1.5.4 Phylogenetic Origin of Truncated tRNA Genes in Cecidomyiids.......155
S1.6 Acknowledgements 156
S1.7 Literature Cited...157

List of Figures

Figure 1 Equilibrium predictions of island biogeographic theory to the relationship between insects and their host plants 32

Figure 2 Predicted effect of plant insularity (genetic distinctiveness) on insect species richness 33

Figure 3 Idealized examples of predicted effects of insularity and ecological opportunity on diversity of phytophagous insects 34

Figure 4 Relationship between plant insularity (x axis), plant species richness (y axis), and gall midge diversity (vertical axis) 35

Figure 5 Relationship between plant species richness (y axis), plant family age (x axis) and gall midge diversity (vertical axis) 36

Figure 6 Relationship between gall midge diversity and plant habitat complexity 37

Figure 7 *Rhopalomyia* phylogenies for A. Cytochrome oxidase subunit I (COI) gene including all North American and Asian species for which data are available. B. Wingless (WG) gene for samples collected within North America. 59

Figure 8 Combined COI and WG phylogeny for *Rhopalomyia* gall midge species 60

Figure 9 Phylogeny of *Asphondylia auripila* group and outgroups 93

Figure 10 Phylogeny of *Asphondylia auripila group* based on combined dataset with ancestral gall position reconstruction by Bayesian methods 94

Figure 11 Phylogenetically independent contrast values for *Asphondylia* ovipositor lengths calculated for each node plotted versus an index of the probability of change in host-plant part usage from one node to the next. 95

Figure 12 Maximum Likelihood reconstructions of ancestral locations for the wingless gene (Wg) and cytochrome oxidase subunit I (COI) and minimum spanning polygons of ancient alleles for the COI locus 121

Figure 13 The marginal posterior probability distributions for between species migration rates estimated with IMa. 122

Figure 14 The marginal posterior probability distributions for ancestral population size estimated from the wingless gene for both species pairs using IMa 123

Figure 15 Predicted probability of the divergence of a lineage of phytophagous insects in different contexts. 134

Figure 16 Organization of the mitochondrial genomes of representatives of the families Cecidomyiidae and Sciaridae 163

Figure 17 Examples of inferred tRNA secondary structure from each cecidomyiid genome. 164

Figure 18 Phylogeny of cecidomyiid genera and outgroup based on sequence data from cytochrome oxidase subunit I gene. 165

Figure 19 All tRNA genes found within the mitochondrial genome of *Mayetiola destructor*. 166

Figure 20 All tRNA genes found within the mitochondrial genome of *Rhopalomyia pomum.* 167

Figure 21 tRNA genes found within the mitochondrial genome of *Asphondylia rosetta* and *Asteromyia spp.* 168

List of Tables

Table 1 Factors affecting diversity under the island phytophagy model, in comparison to similar processes in island biogeographic model....28

Table 2 Multivariate Generalized Linear Model results for both plant families and genera of plant family Asteraceae....29

Table 3 Species richness of Nearctic gall-inducing midges (Diptera: Cecidomyiidae) and associations with species richness, insularity, age, and habitat complexity of plant families....30

Table 4 Species richness of Nearctic gall-inducing midges (Diptera: Cecidomyiidae) and associations with species richness, insularity, and habitat maximum height of plant genera within plant family Asteraceae....31

Table 5 Details of sequence data for each species for each gene used in this study and sampling locations....56

Table 6 Accession numbers for *Rhopalomyia* sequences obtained from Genbank....57

Table 7 *Rhopalomyia* gall morphological characteristics and adult emergence phenology on *Artemisia tridentata*....58

Table 8 Number of sequences obtained per species per gene (see also Table 9)....89

Table 9 Summary of support for monophyly of clades within the *A. auripila* group....90

Table 10 Collection locations for *Asphondylia* samples used in this study. n/a refers to coordinates not available....91

Table 11 Summary of nucleotide variation....118

Table 12 Pairwise comparison of maximum-likelihood estimates (MLE) for reconstructions of ancestral locations of each species in each pair....119

Table 13 Maximum-likelihood estimates (MLE) and 90% highest posterior density (HPD) intervals of IMa model parameters....120

Table 14 Characteristics of protein coding genes of cecidomyiids and other Diptera....162

Chapter 1: General introduction to adaptive radiation and diversification

Adaptive radiation, the rapid evolution of diversity in a lineage associated with different adaptive zones, is responsible for much of the ecological and phenotypic diversity of life and is therefore a central focus of evolutionary biology (Schluter 2000). The processes of natural selection, gene flow, and genetic drift shape both the formation of new species and the adaptation of proliferating species to novel environments during adaptive radiation (Schluter 2000). Thus, different evolutionary processes may play important roles in divergence in different environmental contexts, and combining studies of diversity at multiple levels (above the species level, between species, and within species) can yield strong inferences about the relative roles of the processes involved in divergence.

Phytophagous insects comprise ~25% of terrestrial biodiversity; this remarkable diversity is thought to have arisen through adaptive radiation concomitant with the opportunities provided by the diversification of plant lineages (Ehrlich and Raven 1964; Mitter et al. 1988; Farrell et al. 1991; Thompson 1994; Farrell 1998; Schluter 2000). Ehrlich and Raven's (1964) hypothesis posits that the diversification of plant lineages generates novel adaptive zones for plant-feeding insects and that colonization of these new adaptive zones then leads to adaptive radiation. Thus, adaptive radiation of insect herbivores in these new adaptive zones may involve two main contexts and their combination: 1) shifting between new host-plant species within a higher plant taxon; and

2) adaptation within a single host-plant species to feeding upon different plant resources. Shifts among host-plant species has been considered the most important driver of phytophagous insect diversification, however the importance of shifts among resources on the same host plant has only recently begun to be understood.

1.1 Adaptive Radiation Among Host-plant Species

The idea that adaptive radiation of phytophagous insects proceeds through shifts to novel host plants is an old one (Walsh 1864). The majority of studies have shown that shifts among novel plant taxa provide a major contribution to the origin of new species of phytophagous insects (Via 2001; Berlocher and Feder 2002; Rundle and Nosil 2005; McLeish et al. 2007). The diversification rate of a phytophagous insect lineage should be correlated with the diversity of ecological opportunities available to it (Farrell et al. 2001; McKenna et al. 2009). The bulk of evidence from Coleoptera (McKenna et al. 2009) and Lepidoptera (Wiegmann et al. 2000) shows that rates of diversification are accelerated in Angiosperm feeding lineages relative to sister taxa feeding upon less diverse Gymnosperm relatives (Coyne and Orr 2004). Thus, shifts among host-plant species are an important driver of diversification in phytophagous insects.

1.2 Adaptive Radiation Within Host-plant Species

Following shifts to new host-plant lineages, phytophagous insects may radiate further within the new hosts. Phytophagous insects are generally narrowly specialized on particular host-plant parts or resources (Jaenike 1990; Futuyma 1991). However, novel host-plant species may present the colonizing insect lineages with opportunities in the form of unfilled niches on other plant parts (leaves, stems, roots, flowers, seeds, nodes, or

fruits), unexploited time periods of plant growth (e.g. spring, summer, fall, winter), or insect lineages may diversify their feeding mode on the same plant part (e.g. leaf - folivore, miner, roller, galler). Clades of gall wasps, thrips, and fruit flies have been shown to radiate through utilization of different plant parts (Cook et al. 2002, Condon et al. 2008, and Crespi et al. 2004). Després et al. (2002) showed that globeflower flies have speciated in the absence of host shifts through colonization of different temporal periods of plant resource availability. During adaptive radiation on willow (*Salix* sp.), tenthredinid sawflies have evolved diversity in feeding modes from free feeding ancestors, to leaf fold/rolls, to enclosed leaf galls, stem galls, finally to multilarval stem galls (Nyman et al. 2000). Reproductive isolation during within host adaptive radiation is likely facilitated when shifts to novel plant parts are also accompanied by divergence in life history timing. Thus, divergence in plant-part use increases the strength of divergent selection relative to the same part, and divergence in life history timing may function to prevent or reduce gene flow.

1.3 Summary of Thesis Chapters

The majority of phylogenetic and fossil evidence supports the hypothesis that at higher taxonomic levels, plants and plant-feeding insects have diversified concomitantly (Farrell and Mitter 1998). However, despite many studies of the interactions between insects and their host plants, the mechanisms through which phytophagous insects have radiated adaptively needs to be brought into sharper focus by further studies at both macro- and micro-evolutionary scales.

The goal of this thesis is to bring the mechanisms by which adaptive radiation proceeds, in particular amongst phytophagous insects, into shaper focus using host

specific gall-inducing flies in the family Cecidomyiidae as a model. The thesis contains four main components 1) macroevolutionary determinants of diversity in cecidomyiid flies; 2) host-specific adaptive radiations of cecidomyiid flies; 3) population genetic analyses within a radiation of cecidomyiid flies on a single-host plant species; and 4) molecular and phylogenetic analyses of cecidomyiid mitochondrial genomes. The ordering of the thesis in this fashion allows for a logical transition from macroevolutionary to microevolutionary scales.

First, chapter 2 examines the macro-evolutionary determinants of diversity in cecidomyiid flies. To accomplish this, I apply island biogeographic theory to the diversity of gall inducing midges between host-plant families, and within plant family Asteraceae, to determine the effects of island biogeographic processes on the diversity of gall-inducing midges. The results show that the relationship between diversity of gall-inducing flies and their host plants meets several expectations from island biogeographic theory. Plant taxon genetic distinctiveness, species richness, lineage age, and structural complexity are all found to have significant predictive power in the explanation of variance in the diversity of gall-inducing flies among host-plant species.

Chapter 3 comprises a phylogenetic study of the *Rhopalomyia* gall midges which inhabit host plants within a very large, non-genetically distinctive plant family (Asteraceae). Both host-plant shifts and within-plant speciation events are shown to be important in the diversification of North American *Rhopalomyia.*

In chapter 4, I conduct a phylogenetic and morphometric study of cecidomyiid flies that inhabit creosote bush (*Larrea tridentata*), an abundant, widely distributed, genetically distinct plant in the southwestern deserts of North America. The results

indicate that this group of flies has radiated adaptively within host. I also develop a novel method for conducting independent contrast analyses, which reveals that shifts between parts of the host plant are associated with accelerated evolution in *Asphondylia* ovipositor morphology relative to wing length, a character indicative of overall increase in body size.

In chapter 5, I conduct a population genetic study of divergence in two species pairs of midges that diverged in two different ecological/life-history contexts. The first pair (*Asphondylia auripila* and *Asphondylia foliosa*) both induce galls on the same part of the host-plant species (stem) and display divergence in life-history timing. The second pair (*Asphondylia rosetta* and *Asphondylia florea*) display divergence in plant-part use but adults of these species are temporally coincident. The analyses support a hypothesis of ancient gene flow and large population size during divergence of the species pair that diverged in life-history timing, but no gene flow and small historical population size between the pair that diverged in plant-part use. These results suggest that local adaptation and population divergence are prevented by even low levels of gene flow in small populations, whereas local adaptation in populations of relatively larger effective size may be tolerant to low levels of gene flow during population divergence, a pattern consistent with the neutral theory of molecular evolution and theories of ecological speciation.

In Appendix 1, I examine the evolution of mitochondrial genomes of five cecidomyiid species in four genera. The results show cecidomyiid mitochondrial (mt) genomes to possess four very unusual features. First, cecidomyiid mt genomes are much reduced in size relative to other dipteran mt genomes; second, cecidomyiid mt genomes

display elevated A + T content - more than 83% of the coding region; third, cecidomyiid mt genomes display rearrangement of tRNA genes; fourth, the most unusual feature of cecidomyiid genomes examined in this study is an extreme truncation of all tRNA genes, including the loss of TC arms and apparent absence of the 3' part of the aminoacyl stems. The truncated and rearranged tRNA genes shown in this study illustrate the dynamic nature of cecidomyiid mitochondrial genomes and extend the taxonomic breadth of the observation that in some lineages tRNA genes are severely truncated to Diptera.

1.4 Literature Cited

Berlocher, S.H., J.L. Feder. 2002. Sympatric speciation in phytophagous insects: Moving beyond controversy? Ann. Rev. Entomol. 47: 773–815.

Cook, J.M., A. Rokas, M. Pagel, and G.N. Stone. 2002. Evolutionary shifts between host oak sections and host plant organs in *Andricus* gall wasps. Evolution 56: 1821-1830.

Condon, M.A., S.J. Scheffer, M.L. Lewis, and S.M. Swensen. 2008. Hidden neotropical diversity: greater than the sum of its parts. Science 320:928-931.

Coyne, J. and A. Orr. 2004. *Speciation*. Sinaeur Associates, Inc. Sunderland, MA.

Crespi, B. J., D.C. Morris, and L.A. Mound. 2004. Evolution of Ecological and Behavioural Diversity: Australian Acacia Thrips as Model Organisms. Australian Biological Resources Study & CSIRO Entomology: Canberra, Australia, 328 pages.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. 2002. Speciation in the globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host

plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258–268.

Ehrlich, P.R. and P.H. Raven. 1964. Butterflies and plants: a study in co-evolution. Evolution 18:586-608.

Farrell, B.D. 1998. "Inordinate fondness" explained: why are there so many beetles? Science 281:555-559.

Farrell, B.D., D. Dussourd, and C. Mitter. 1991. Escalation of plant defense: do latex/resin canals spur plant diversification? Am. Nat. 138:881-900.

Farrell, B.D., A.S. Sequeira, B.C. O'Meara, B.B. Normark, J.H. Chung, and B.H. Jordal. 2001. The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). Evolution 55:2011-2027.

Futuyma, D.J. 1991. Evolution of host specificity in herbivorous insects: genetic, ecological, and phylogenetic aspects, pp. 431-454. In P.W. Price, T.M. Lewinsohn, G.W. Fernandes, and W.W. Benson (eds), Plant-animal interactions: evolutionary ecology in tropical and teperate regions. John Wiley and Sons, New York.

Jaenike, J. 1990. Host specialization in phytophagous insects. Annu. Rev. Ecol. Syst. 21:243-273.

McKenna, D.D., A. Sequeira, A.E. Marvaldi, and B.D. Farrell. 2009. Temporal lags and overlap in the diversification of weevils and flowering plants. PNAS 106:7083-7088.

McLeish, M. J., B.J. Crespi, T.W. Chapman, and M.P. Schwarz. 2007. Parallel diversification of Australian gall-thrips on *Acacia*. Mol. Phylogenet. and Evol. 43:714-725.

Mitter, C., B. Farrell, and B. Wiegmann. 1988. The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? Am. Nat. 132:107-128.

Nyman, T., A. Widmer, and H. Roininen. 2000. Evolution of gall morphology and host-plant relationships in willow-feeding sawflies (Hymenoptera: Tenthredinidae). Evolution 54: 526-533.

Rundle, H.D., and P. Nosil. 2005. Ecological speciation. Ecol. Lett. 8:336-352.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford, Oxford University Press.

Thompson, 1994. The coevolutionary process. University of Chicago Press, Chicago.

Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends Ecol. Evol. 16 381-390.

Walsh BD. 1864. On phytophagic varieties and phytophagic species. Proc. Entomol. Soc. Phil. 3:403–30

Wiegmann, B.M., C. Mitter, J.C. Regier, T.P. Friedlander, D.M. Wagner, and E.S. Nielsen. 2000. Nuclear genes resolve Mesozoic-aged divergences in the insect order Lepidoptera. Mol. Phylogent. Evol. 15:242-259.

Chapter 2: Island phytophagy

2.1 Abstract

Island biogeographic theory has been applied to spatially discrete habitats such as oceanic islands (Losos and Schluter 2000; Parent and Crespi 2006), mountain tops (Whittaker and Fernandes-Palacios 2007) and lakes (Whittaker and Fernandes-Palacios 2007) but not to habitats where distance is defined by genetic differences. Plants, to the insects feeding upon them, present notable parallels with oceanic islands and island-like habitats (Janzen 1968, Strong et al. 1984). Like islands, plant taxa vary in insularity (genetic distance between plants), area (abundance and distribution; Strong et al. 1984), habitat diversity (plant structural complexity; Lawton and Schroeder 1977), and island age (lineage age; Fernandes 1992); moreover speciation may occur on the host plant (Joy and Crespi 2007; Condon et al. 2008), as within islands (Losos and Schluter 2000; Whittaker and Fernandes-Palacios 2007), or between host species (between islands, Dres and Mallet 2002).

Plant-feeding insects have undergone unparalleled diversification relative to non-plant feeding relatives (Ehrlich and Raven 1964; Farrell 1998). Some evidence has supported a role for coevolution of insects with plants (Ehrlich and Raven 1964; Farrell 1998) and phylogenetic constraint (Price 1994; Leyva et al. 2003) in the evolution of the diversity of phytophagous insects. Yet on balance the underlying causes of phytophagous insect diversity remain to be discerned and quantified (Janz et al. 2006). Here I show that the relationship between diversity of gall-inducing flies and their host plants meets

several expectations from island biogeographic theory. Plant taxon insularity, species richness, age, and structural complexity are all found to have significant predictive power in the explanation of variance in the diversity of gall-inducing flies among host-plant species. The application of island biogeographic theory to insects and their host plants provides a novel, tractable framework for integrating diverse causal factors to explain the remarkable diversity of plant feeding insects.

Keywords: Macroevolution, insularity, lineage age, ecological opportunity, adaptive radiation, Cecidomyiidae

2.2 Introduction

Island biogeographic theory (MacArthur and Wilson 1967) posits that diversity on oceanic islands, or a suitable patch of habitat surrounded by unsuitable habitat, is determined by equilibrium between colonization, emigration, and extinction. Island diversity equilibrium points are influenced by the size of the island, distance of the island from the source population, the age of the island, and the diversity of habitats present on the island (MacArthur and Wilson 1967; Simberloff 1976). To host-specific insects, host plant species are like islands surrounded by unsuitable habitat (inhospitable host plants). Thus, as with oceanic islands, the diversity of phytophagous insects on plants may represent an equilibrium between colonization, emigration and extinction (Figure 1). In theory, the equilibrium point for the diversity of insects on a given plant species or higher plant taxon should be affected by the abundance of the host plant species, plant phylogenetic distinctiveness, the number of plant species in the lineage, the age of the plant lineage, and the complexity of the habitat provided by the plant, as shown in Figures 2 and 3, and Table 1. This conceptual framework provides clear, testable

predictions regarding the causes of variation in species diversity among plant-feeding insects, in relation to characteristics of their hosts.

An ideal opportunity for testing and evaluating the predictions of island biogeographic theory in the context of the diversity of phytophagous insects is provided by the more than 1200 species of gall-inducing midges (Diptera: Cecidomyiidae) in North America. Gall midges are highly host-plant specific, disperse easily over vast distances (Mamaev and Krivosheina 1993) and have radiated within a variety of host-plant species (Fernandes 1992). The species richness of gall-inducing insects on host-plant species is positively associated with number of species in the plant taxon (Fernandes 1992; Roskam 1985) and both colonization and within plant speciation (Joy and Crespi 2007; Stireman et al. 2008) have been shown to be important in their diversification.

I analyzed the known host preferences of gall-inducing midges (Diptera: Cecidomyiidae) on host-plant species within North America together with a plant family-level phylogeny (Davies et al. 2004), and a phylogeny for plant family Asteraceae (Funk et al. 2005), to test how well island biogeographic theory predicts variation in the diversity of gall-inducing midges. Specifically, I tested four hypotheses stemming from island biogeographic theory: (1) distantly-related ('more insular') host-plant species should have lower species diversity than closely-related plant species; (2) as the number of species in a plant family or genus (ecological opportunity) increases, so should the number of host-specific gall-inducing insects; (3) older plant lineages should host a greater proportion of gall-inducing midges than younger taxa; and (4) plants which are structurally more complex (have more foliar area, more plant parts, more temporal

variation in availability of plant parts), should harbour a larger gall-inducing midge fauna.

2.3 Methods

2.3.1 Host-plant preferences database

The host-plant preference database consisted of detailed host-plant preferences for 758 species from 67 genera of gall-inducing midge species on 71 plant families within North America and 213 species from 18 genera of gall-inducing midge species on 61 plant genera within the North American Asteraceae. These data were obtained from published sources excluding unknown species and species for which the host plant is undetermined (Gagné 1989). Cecidomyiidae is a very large family of flies and as such there are species not yet described or recorded; however, sampling effort has been greatest in North America (Dorchin 2008) and missing species are unlikely to be biased towards particular plant taxa in a systematic fashion. A detailed table showing the number of gall-inducing midge species by plant family, and by plant genus within Asteraceae, can be found in Table 3 and Table 4.

2.3.2 Statistical methods

Count data are well-modeled using a Poisson distribution, but this approach typically needs to be modified to account for larger variances in the counts than can be modeled using simple likelihood methods (McCullagh and Nelder 1989). This is routinely done using quasilikihood methods that permit an additional parameter to account for overdispersion in the counts (McCullagh and Nelder 1989). For analyses at the plant family level and within plant family Asteraceae, I employed generalized linear

models (GLM) with quasipoisson distributions specifying an extra parameter to account for the inherent overdispersion in the counts of the numbers of gall-inducing midge species. GLM analyses included cecidomyiid species richness as dependent variables and plant species richness, plant insularity, plant lineage age, and habitat variables as independent variables. Variables were not significantly intercorrelated.

2.3.3 Estimation of plant species richness diversity relationships

I used published data on the sizes of North American plant families (Table 3) and the size of each genus of Asteraceae (Table 4). I used the number of species in each plant taxon as a surrogate for ecological opportunity (Parent and Crespi 2006). I used these plant family size values with gall-inducing midge species richness in multivariate generalized linear models with plant insularity and plant lineage age to control for the influence of other factors.

2.3.4 Inference of plant insularity and insect diversity

I calculated insularity for each plant family using a phylogeny of angiosperms (Davies et al. 2004), and for each asterid genus using a phylogeny of major clades of Asteraceae (Funk et al. 2005). I employed the Vane-Wright (1991) measure of taxonomic distinctiveness to calculate insularity values for each taxon on each phylogenetic tree. To control for the effects of plant taxon size, these insularity values were then used in multiple regression analyses.

2.3.5 Plant taxon age

To test the hypothesis that older plant families would host greater numbers of gall-inducing midges relative to younger taxa, cecidomyiid species richness was

compared to plant family ages determined with a combination of molecular data calibrated with available fossil data (Wikstrom et al. 2004; Magallon and Castillo 2009; Table 3). Fossil data were not available for Asteraceae and the phylogeny lacked appropriate branch lengths; thus, tests of the taxon age hypothesis were not feasible among genera of Asteraceae.

2.3.6 Plant complexity and insect species richness

To test the hypothesis that plant-structural complexity (niche space and habitat diversity) also contribute to explaining the variance in diversity of phytophagous insects between plant taxa, I compared gall-inducing midge species richness (1) at the family level with plant attributes (for example woody versus herbaceous); (2) within Asteraceae with plant maximum height. For these analyses plant height and plant attributes are considered surrogates for habitat complexity.

2.4 Results

To test hypothesis (1) that plant taxa with higher insularity (plants that are more genetically and taxonomically distinct) are expected to have lower diversity of gall-inducing insects, I compared the number of species of gall-inducing midges with plant insularity values calculated from phylogenetic trees at the plant family (Davies et al. 2004) level and for family Asteraceae (Funk et al. 2005). Insularity values were calculated for each plant taxon using the Vane-Wright method (Vane-Wright 1991), whereby each plant taxon is assigned the inverse value of the number of nodes connecting it to the root of the plant phylogeny. Thus, plant taxa displaying relatively higher insularity values are more genetically distinctive. Comparisons between dependent

and independent variables were modeled using Generalized Linear Models (GLM) in which the number of cecidomyiid species inhabiting a plant taxon was the dependent variable and plant insularity, plant species richness, and plant lineage age were independent variables (Table 2). Figure 4 reveals that, both among plant families and among genera within family Asteraceae, insularity is a significant predictor of gall-inducing midge species richness. Tests of hypothesis (1) thus illustrate that both among plant families, and among genera within family Asteraceae, plant taxa of lower insularity harbour relatively large numbers of cecidomyiid species (Tables 3 and 4).

To test hypothesis (2) that as the number of species in the plant taxon (a proxy for long-term ecological opportunity) increases, the diversity of cecidomyiid species should also increase, cecidomyiid species richness was compared with plant taxon richness using the same model framework (Table 2). Figure 4 illustrates that in North America, the number of species in the plant taxon explains a significant proportion of between plant taxon variance in the number of species of gall-inducing midges, both among plant families, and among plant genera within the family Asteraceae.

To test hypothesis 3 that plant lineage age should be positively associated with increases in cecidomyiid species richness, plant family age (as determined from molecular dating utilizing available fossils as calibration points (Wikstrom et al. 2004; Magallon and Castillo 2009)), was compared with cecidomyiid species richness. After controlling for plant taxon species richness and plant family insularity, the age of plant families is positively associated with cecidomyiid species richness (Table 2, Figure 5).

Hypothesis 4 predicts that plant species with more niche space will show elevated species richness of plant-feeding insects relative to plant species with fewer niches.

Figure 6 and Table 2 illustrate that plant taxa which are woody and plant taxa which are taller (proxies for plant architectural diversity) show higher diversity of gall-inducing midge species.

2.5 Discussion

Taken together, the result of testing these four predictions show that in gall-inducing cecidomyiid dipterans, insect species diversity can be predicted from a combination of host plant insularity, species richness, age, and habitat diversity. How well does this predictive, integrative conceptual framework help to explain insect diversity among other insect-plant associations?

2.5.1 Insularity

Less-insular plant taxa may harbour higher diversity of insects in large part because shifts between host-plant species are one of the most important drivers of diversification in phytophagous insects (Abrahamson et al. 1994; Janz & Nylin 1998; Dobler &Farrell 1999; Ronquist & Liljeblad 2001; Berlocher & Feder 2002), and host plant shifts are more common between plant taxa that are more closely related (Farrell 1998; Nyman et al. 2006). Closely-related host plants are more likely to be ecologically similar relative to distantly related hosts, facilitating colonization of, and adaptation to, the new host. Consistent with this idea, studies of sawflies (Hymenoptera: Tenthredinoidea) have shown that shifts between host plants which are closely related are more common than those between plants which are distantly related, moreover, when shifts between distantly related plants occur, the plants are often ecologically similar (Price 1994; Nyman et al. 2006). Similarly, shifts between host plants by Australian gall-

inducing thrips (Thysanoptera: Phlaeothripidae) take place more often between closely-related *Acacia* species than distantly-related ones, and in rare instances when thrips do shift to distantly related *Acacia* species, such shifts are accompanied by dramatic phenotypic and life history changes in the colonizing thrips lineage (Crespi et al. 2004; McLeish et al. 2007). Diversification of phytophagous beetle species (Coleoptera) has similarly been shown to reflect a complex history of coevolution between beetles and angiosperms in which host plant conservatism is observed (Farrell 1998; McKenna et al. 2009).

Host shifts may facilitate diversification in part through specialization on novel host plants. By this process, specialization may proceed through either preference and performance tradeoffs between hosts, or via ecological divergence of host-shifted populations. In each case, selection favours the evolution of reproductive isolation such that host races perform better on their own host (Joshi & Thompson 1997; Feder et al. 1999; Hawthorne et al. 2001; Nosil et al. 2002).

Thus, host-plant taxon genetic distinctiveness may be an important factor in the diversification of phytophagous insects by strongly influencing the success of insect attempts to colonize a plant taxon.

2.5.2 Ecological Opportunity (Plant Diversity)

Much recent and historical work supports a key role for plant diversity in explaining the diversity of phytophagous insects (Novotny et al. 2006). Previous studies of cecidomyiid flies in the Indonesian islands have shown that the most important predictor of cecidomyiid species richness was the diversity of the plant community

(Fernandes 1992). Similarly, studies of cecidomyiid flies in the Palearctic have shown that plant species richness is an important predictor of insect diversity (Roskam 1985). Novotny et al. (2006) showed that the disparity in insect species richness between temperate and tropical areas is primarily a function of the increased plant species richness found in tropical areas. As the number of plant species increases, so too may the number of ecological opportunities for phytophagous insects in the form of more species to colonize, more diversity in physical plant parts (Condon et al. 2008), higher variance in duration of plant parts (Després et al. 2002), and greater phenological variance in the distribution of plant growth periods (Joy and Crespi 2007).

2.5.3 Plant Lineage Age

Provided niche space remains available, species richness of phytophagous insects on host plants should increase proportionally with the length of time hosts and colonizing insects are in contact (Strong et al. 1984; White et al. 2006; Brandle et al. 2008). Following colonization of a new habitat most niches on the plant are likely to be empty. Through time open niches on the plant would gradually become filled through colonization and speciation. As the plant colonizes and adapts to new environments novel unfilled niches may open up to phytophaous insects in the form of novel plant parts, or novel time periods of plant growth. Further some colonizing insects may compete more effectively for niche space in different environments increasing the diversity of phytophagous insects across the range of the host plant.

Previous studies of phytophagous insects in relation to plant lineage age have shown that on woody host species, the number of species of Lepidoptera and Auchenorrhyncha increased over time (Brandle et al. 2006). Farrell and Mitter (1994)

showed that in accordance with predictions from Erhlich and Raven (1967), much of diversity of phytophagous insects can be attributed to effects of plant lineage age and ecological opportunity. In contrast, other studies have not found positive associations between diversity and increases in island age (Parent and Crespi 2006, on snails) or plant lineage age (Roskam 1985; Fernandes 1992). The niche breadth of phytophagous insects may, in part, determine the likelihood of a relationship with host plant lineage age, for any particular insect taxon. For example, more-generalist plant-feeding insects may colonize hosts regardless of age, while more specialized plant-feeding insects may display tighter phylogenetic associations with particular plant lineages, given their closer degree of adaptation to particular plant lineages. Thus, relationships between plant lineage age and species richness of more specialized plant-feeding insects may be strongly influenced by rare switches between disparate plant lineages.

2.5.4 Plant Architectural Complexity

Plant architectural diversity likely promotes insect diversification through increases in available niche space (Condon et al. 2008). Whereby, plants that are more architecturally diverse exhibit more variation in types of plant parts (Cuevas-Reyes 2004), which in turn allows extreme specialization such that the result is the fine division of a host plant into discrete niches associated with different plant parts (Cook et al. 2002; Després et al. 2002; Condon et al. 2008). Condon and Steck (1997) showed that *Blephanoneura* fruit flies specialized on either male or female flowers within the same species of Guraniinae, they suggest that adaptation to the different-sex flower may involve relatively few changes in adult or larval physiology.

2.5.5 Plant Abundance and Distribution

An additional important factor governing diversity of phytophagous insects, which could not be accounted for in the analyses conducted here, is the role of host-plant geographic range size, the equivalent of island size in island biogeographic theory. Host-plant taxa which have larger geographic ranges may facilitate diversification of phytophagous insects through geographical isolation of the insect taxa in various parts of the host range. Host-plant taxa with large geographic ranges often span a range of environmental conditions, and geographically-associated genetic variation in host plants may facilitate speciation in insect herbivores via reinforcement despite insect gene flow across the host plant range, because insect performance can be tightly coupled with host plant genetic variation. Associations between host-plant geographic range and insect species diversity have been demonstrated for many insect groups (Fowler and Lawton 1982; Blanche and Westoby 1996; Kelly and Southwood 1999), and geographic distribution of the host-plant species was found to be an especially-important driver of diversity of neotropical fruit flies in the genus *Blepharoneura* (Condon et al. 2008). Crespi et al. (2004, page 57) likewise found a positive association between the geographic range size of *Acacia* species and diversity of phyllode-gluing thrips species in Australia.

As with oceanic islands where larger islands derive a greater proportion of speciation events from within island speciation (Losos and Schluter 2000), host plant taxa with larger geographic ranges and higher abundance within that range should be expected to experience more within-plant speciation (Joy and Crespi 2007; Condon et al. 2008) relative to host plants with smaller ranges and those which are less dense within their

range. Consistent with this idea, Crespi et al. (2004, page 57) showed that *Acacia* species with the largest geographic ranges have a higher proportion of phyllode-gluing thrips species that are specialized on a single *Acacia* species. Similarly, gall-inducing midges which inhabit plants with large geographic ranges and higher abundance within those ranges also appear to have a large number of closely related species living on them - for example, *Asphondylia* gall midges which inhabit creosote bush include 15 species, and *Rhopalomyia* gall midges which inhabit big sagebrush include 26 species, many of which appear to be derived from within host speciation events.

Previous studies of *Blephanoneura* fruit flies (Condon et al. 2008) and cecidomyiid flies (Joy and Crespi 2007; Stireman et al. 2008), have shown that diversification may often be associated with specialization on different host-plant resources in the absence of shifts to a novel host-plant species. Studies of diversification in oak gall wasps (Cook et al. 2002) have also implicated shifts to new organs on the same host-plant associated with divergence without a shift to a new host plant. Within host-plant speciation has also been inferred as a plausible mechanism of diversification for Australian thrips which induce galls on *Acacia aneura* (Crespi et al. 2004). Future tests of island biogeographic theory in this context may be able to identify, as with oceanic islands (Losos and Schluter 2000), if there is a threshold range size and abundance beyond which speciation events are predominantly within host relative to between host.

2.5.6 Conclusion

Island biogeographic theory (MacArthur & Wilson 1967) has proven to be a remarkably useful framework for understanding what causal factors shape diversity on

oceanic islands and other contexts of insularity such as mountaintops and isolated lakes. Experimental tests, in which arthropods were removed from oceanic islands and then recolonized, have shown the main tenets of the theory to be largely experimentally validated (Simberloff 1976). By the conceptual framework described here, island biogeographic theory also provides a useful framework for integrating the evolutionary processes that have determined the remarkable diversity of phytophagous insects. The analyses conducted here represent the first comprehensive test of island biogeographic theory as applied to the diversity of plant-feeding insects (Janzen 1968; Strong et al. 1984) and indeed the first test beyond spatial contexts of insularity such as mountaintops and mid-oceanic islands.

2.6 Acknowledgments

This work was supported by an NSERC operating grants to Bernard J. Crespi and Arne O. Mooers. I especially thank Tanja Schwander, Arne Mooers, Will Stein, and Laura Weir for helpful discussion and encouragement. Ruth Joy and Pete Hurd provided invaluable statistical advice. I am also grateful to the Simon Fraser University FAB* Lab.

2.7 Literature Cited

Abrahamson, W.G., J.M. Brown, S.K. Roth, D.V. Summerford, J.D. Horner, M.D. Hess, S.T. How, T.P. Craig, R.A. Packer, and J.K. Itami. 1994. Gallmaker speciation: an assessment of the roles of host-plant characters, phenology, gallmaker competition, and natural enemies. pp. 208–222 *in* Price PW, W.J. Mattson and Y.N. Baranchicov. (eds) *The Ecology and Evolution of Gall-Forming Insects.* US Department of Agriculture.

Berlocher, S.H. and J.L. Feder. 2002. Sympatric speciation in phytophagous insects: Moving beyond controversy? Ann. Rev. Entomol. 47: 773–815.

Blanche K.R. and M. Westoby. 1996.The effect of the taxon and geographic range size of host eucalypt species on the species richness of gall-forming insects. Austr. J. Ecol. 21: 332–335.

Brandle, M., I. Kuhn, S. Klotz, C. Belle, and R. Brandl. 2008. Species richness of herbivores on exotic host plants increases with time since introduction of the host. Biodiv. Res. 14:905-912.

Condon, M.A., S.J. Scheffer, M.L. Lewis, and S.M. Swensen. 2008. Hidden neotropical diversity: greater than the sum of its parts. Science 320:928-931.

Cook, J.M., A. Rokas, M. Pagel, and G.N. Stone. 2002. Evolutionary shifts between host oak sections and host plant organs in *Andricus* gall wasps. Evolution 56: 1821-1830.

Crespi, B. J., D.C. Morris, and L.A. Mound. 2004. *Evolution of Ecological and Behavioural Diversity: Australian Acacia Thrips as Model Organisms*. Australian Biological Resources Study & CSIRO Entomology: Canberra, Australia, 328 pages.

Cuevas-Reyes, P., M. Quesada, P. Hanson, R. Diroz, and K. Oyama. 2004. Diversity of gall-inducing insects in a Mexican tropical dry forest: the importance of plant species richness, life-forms, host plant age and plant density. J. Ecol. 92:707-716.

Davies, T.J., T.G. Barraclough, M.W. Chase, P.S. Soltis, D.E. Soltis, and V. Savolainen. 2004. Darwin's abominable mystery: insights from a supertree of the angiosperms PNAS 101: 1904-1909.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. 2002. Speciation in the globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258–268.

Dorchin, N. 2008. Gall midges (Diptera: Cecidomyiidae). pp. 1576-1580 *in* J.L. Capinera (ed) *Encyclopedia of Entomology*, Volume 4. Springer.

Dobler, S. and B.D. Farrell. 1999. Host use evolution in *Chrysochus* milkweed beetles: evidence from behavior, population genetics and phylogeny. Mol. Ecol. 8: 1297–1307.

Ehrlich, P.R. and P.H. Raven. 1964. Butterflies and plants: a study in co-evolution. Evolution 18:586-608.

Farrell, B.D. and C. Mitter. 1994. Adaptive radiation in Insects and plants: time and opportunity. Amer. Zool. 34:57-69.

Farrell, B.D. 1998. "Inordinate fondness" explained: why are there so many beetles? Science 281:555-559.

Feder, J. L. and K.E. Filchak. 1999. It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. Entomol. Exp. Appl. 91:211–225.

Fernandes, G.W. 1992. Plant family size and age effects on insular gall-forming species richness. Glob. Ecol. and Biogeog. Lett.

Fowler, S.V. and J.H. Lawton. 1982. The effects of host plant distribution and local abundance on the species richness of agromyzid flies attacking British Umbellifers. Ecol. Entomol. 7: 257–266.

Funk, V. A., R.J. Bayer, S. Keeley, R. Chan, L. Watson, B. Gemeinholzer, E. Schilling, J.L. Panero, B.G. Baldwin, N.T. Garcia-Jacas, A. Susanna, and R.K. Jansen. 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. In: Friis, I. and H. Balslev [eds.], Proceedings of a Symposium on Plant Diversity and Complexity Patterns - Local, Regional and Global Dimensions. The Royal Danish Academy of Sciences and Letters, Copenhagen. Biol. Skr. 55:343-374.

Gagné, R. J. 1989. *The Plant-Feeding Gall Midges of North America*. Cornell University Press, Ithaca.

Hawthorne, D.J. and S. Via. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412: 904–907.

Hunt, T., J. Bergsten, Z. Levkanicova, A. Papadopoulou, O. St. John, R. Wild, P.M. Hammond, D. Ahrens, M. Balke, M.S. Caterino, J. Gomez-Zurita, I. Ribera, T.G. Barraclough, M. Bocakova, L. Bocak, and A.P. Vogler. 2007. A Comprehensive Phylogeny of Beetles Reveals the Evolutionary Origins of a Super radiation. Science 318:1913-1916.

Janz N. and S. Nylin. 1998. Butterflies and plants: A phylogenetic study. Evolution 52: 486–502.

Janz, N., S. Nylin, and N. Wahlberg. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects BMC Evol. Biol. 6: 4.

Janzen, D. 1968. Host plants as islands in evolutionary and contemporary time. Am. Nat. 102, 592-595.

Joshi, A. and J.N. Thompson. 1997. Adaptation and specialization in a two-resource environment in *Drosophila* species. Evolution 51: 846–855.

Joy, J.B. and B.J. Crespi. 2007. Adaptive radiation of gall-inducing insects within a single host-plant species. Evolution 61: 784-795.

Kelly, C.K. and T.R.E. Southwood. 1999. Species richness and resource availability: A phylogenetic analysis of insects associated with trees. *PNAS* 96: 8013–8016.

Lawton, J.H. and D. Schroder. 1977. Effects of plant type, size of geographical range and taxonomic isolation on number of insect species associated with British plants. Nature 265:137-40.

Leyva, K.J., K.M. Clancy, and P.W. Price. 2003. Oviposition strategies employed by the western spruce budworm: tests of predictions from the phylogenetic constraints hypothesis. Agric. For. Entomol. 5: 9-16.

Losos, J.B. and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. Nature 408, 847-850.

MacArthur, R.H. and E.O. Wilson. 1967. *The theory of island biogeography*. Princeton Univ. Press, Princeton.

Magallon, S. and A. Castillo. 2009. Angiosperm diversification through time. Am. J. Bot. 96:349-365.

Mamaev B.M. and N.P. Krivosheina. 1993. *The larvae of the gall midges* (Diptera, Cecidomyiidae). A.A.Balkema (Rotterdam) Brockfield.

McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*. Chapman and Hall, London.

McKenna, D.D., A. Sequeira, A.E. Marvaldi, and B.D. Farrell. 2009. Temporal lags and overlap in the diversification of weevils and flowering plants. PNAS 106:7083-7088.

McLeish, M. J., B.J. Crespi, T.W. Chapman, and M.P. Schwarz, M. P. 2007. Parallel diversification of Australian gall-thrips on *Acacia*. Mol. Phylogenet. and Evol. 43:714-725.

Nosil, P., B.J. Crespi, and C.P. Sandoval. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417: 440–443.

Nosil, P., C.P. Sandoval, and B.J. Crespi. 2006. The evolution of host preference in allopatric vs. parapatric populations of *Timema cristinae* walking-sticks. J. Evol. Biol. 19: 929-942.

Novotny, V., P. Drozd, S.E. Miller, M. Kulfan, M. Janda, Y. Basset, and G.D. Weiblen. 2006. Why are there so many species of herbivorous insects in tropical rainforests? Science 313:1115-1118.

Nyman, T., B.D. Farrell, A.G. Zinovjev, and V. Vikberg. 2006. Larval habits, host-plant associations, and speciation in nematine sawflies (Hymenoptera: Tenthredinidae. Evolution 60:1622-37.

Parent, C.E. and B.J. Crespi. 2006. Sequential colonization and diversification of Galapagos endemic land snail Genus *Bulimulus* (Gastropoda, Stylommatophora). Evolution 60:2311-28.

Price, P.W. 1994. Phylogenetic constraints, adaptive syndromes, and emergent properties: from individuals to population dynamics. Res. Pop. Ecol. 36:3-14.

Ronquist, F. and J. Liljeblad. 2001. Evolution of the gall wasp host plant association. Evolution 55: 2503–2522.

Roskam, J.C. 1985. Evolutionary patterns in gall midge–host plant associations (Diptera: Cecidomyiidae). *Tijdschrift voor Entomologie*, 128: 193–213.

Simberloff, D.S. 1976. Species turnover and equilibrium island biogeography. Science 194:572-578.

Strong, D.R., J.H. Lawton, and T.R.E. Southwood. 1984. *Insects on plants. Community patterns and mechanisms*. Blackwell, London.

Stireman J.O., E.M. Janson, T.G. Carr, H. Devlin, and P. Abbot. 2008. Evolutionary radiation of *Asteromyia carbonifera* (Diptera: Cecidomyiidae) gall morphotypes on the goldenrod *Solidago altissima* (Asteraceae). Biol. J. Linn. Soc. Lond. 95:840-858.

Vane-Wright, R.I., C.J. Humphries, and P.H. Williams. 1991. What to protect? Systematics and the agony of choice. Biol. Conserv. 55: 235-54.

Whittaker, R.J. and J.M. Fernandes-Palacios. 2007. *Island biogeography: ecology, evolution and conservation*. Oxford Univ. Press, Oxford.

White, E.P., P.B. Adler, W.K. Lauenroth, R.A. Gill, D. Greenberg, D.M. Kaufman, D.M., A. Rasswiler, J.A. Rusak, M.D. Smith, J.R. Steinbeck, R.B. Waide, and J. Yao. 2006. A comparison of the species-time relationship across ecosystems and taxonomic groups. Oikos 112:185-195.

Wikstrom, N., V. Savolainen, and M.W. Chase. 2004. Evolution of the angiosperms: Calibrating the family tree. Proc. Roy. Soc. B, 268:2211-2220.

Table 1 Factors affecting diversity under the island phytophagy model, in comparison to similar processes in island biogeographic model. Solid arrows represent direction of effects of island biogeography, and unfilled arrows represent effects of island phytophagy. Diagonal lines indicate no predicted effect.

Island Phytophagy Process	Colonization	Within Plant (Island) Speciation	Extinction	Island Biogeography Process
Plant Insularity	⇩ More insular plants are more difficult for insects to adapt to and colonize	⇧ Insect lineages which have colonized an insular plant lineage will experience reduced competition and will be more likely to speciate within host		
	More isolated islands are more difficult to reach by potential colonists ⬇	Species colonizing a distant island experience reduced competition thus more likely to speciate within island ⬆		**Island Insularity**
Plant Species Richness	⇧ More plant species increases the likelihood that a potential insect colonist will find suitable habitat	⇩ More plant species increase the likelihood that an insect competitor would inhibit within-plant speciation	⇩ More plant species increase ecological opportunity thus increase potential to expand niche in changing conditions	
	More diversity of habitats increasesthe likelihood that a colonist will find habitat ⬆	More habitat diversity on island should increase the likelihood of within island speciation due to more open niches ⬆	More habitat diversity on an island should allow species more latitude to adapt to changing environments ⬇	**Habitat Diversity**
Plant Abundance and Range Size	⇧ As the range size and abundance of a plant taxon increases so does the opportunity for contact with potential colonizing insects	⇧ More opportunity for isolation by distance. More opportunity for insect to locally adapt and speciate via reinforcement despite gene flow	⇩ Insects ihabiting plant taxa with large range sizes are less likely to go extinct because of changing local conditions	
	Larger islands are a larger target for colonizing species thus colonization of larger islands is higher than small ones ⬆	More opportunity for isolation by distance. Larger islands also likely have higher habitat diverstiy which should increase within island speciation ⬆	Species on larger islands may have larger population sizes thereby reducing extinction risk ⬇	**Island Area**
Plant Lineage Age	⇧ Older plant lineages have had more time for insects to colonize. All else being equal old plant taxa should have more species	⇧ Older plant lineages have had more time for speciation events to occur	⇧ Older plant lineages have had more time for their inhabiting insects to go extinct	
	Older islands have had more time for potential colonists to reach them. Thus older islands more species. ⬆	Older islands have had more time for speciation events to occur ⬆	Older islands have had more time for their inhabiting species to go extinct ⬆	**Island Age**
Plant Architectural Diversity	⇧ The more plant parts displayed by a plant taxa the more likely an insect colonist is to find a suitable open niche.	⇧ All else being equal plant taxa which have more types of plant parts have more niches to be filled by within plant speciation		
	Higher elevation islands are a bigger colonizatin target and also have more habitat diversity ⬆	Higher islands are likely to have more habitat diversity and thus more open niches for within island speciation ⬆		**Island Elevation**

Table 2 Multivariate Generalized Linear Model results for both plant families and genera of plant family Asteraceae. The sample size (n), ḅ b S.E., and p-value are provided for each independent variable in each model.

Model and Response Variables	Independent Variables	n	b	b S.E.	P-value
Model: Plant Families					
Cecidomyiid Species Richness	Plant Family Insularity	30	-2.759	0.939	0.006
Cecidomyiid Species Richness	Plant Family Species Richness	30	0.493	0.168	0.002
Cecidomyiid Species Richness	Plant Lineage Age	30	3.832	1.185	0.002
Model: Asteraceae					
Cecidomyiid Species Richness	Plant Genus Insularity	31	-2.721	1.082	0.017
Cecidomyiid Species Richness	Plant Genus Species Richness	31	0.584	0.166	0.001
Cecidomyiid Species Richness	Plant Genus Maximum Height	31	0.811	0.319	0.016

Table 3 Species richness of Nearctic gall-inducing midges (Diptera: Cecidomyiidae) and associations with species richness, insularity, age, and habitat complexity of plant families.

Plant Family	Number of Cecidomyiid Species	Plant Family Species Richness	Plant Family Insularity	Plant Family Age (MYA)	Habitat Complexity
Amaranthaceae	2	2400	0.001761342	34	herb/woody
Apiaceae	3	3000	0.001291651	81.255	herb
Apocynaceae	3	1500	0.001549981	49	herb
Asteraceae	213	23000	0.001291651	91.02	herb/woody
Boraginaceae	1	2000	0.001614564	79	herb
Brassicaceae	8	3700	0.001383912	65.97	woody
Buxaceae	1	90	0.003522684	115.67	herb
Caprifoliaceae	15	800	0.001761342	56	woody
Convolvulaceae	3	1650	0.001549981	65.5	woody
Cornaceae	9	68	0.002039449	101.73	woody
Cucurbitaceae	2	825	0.001383912	76.07	woody
Dioscoreaceae	1	750	0.003229127	115.3	woody
Euphorbiaceae	2	7500	0.001684762	70	herb
Fabaceae	40	19400	0.001614564	76.5	herb/woody
Fagaceae	30	900	0.001549981	93.5	woody
Hamamelidaceae	2	80	0.002421845	60.5	herb
Lamiaceae	18	7200	0.001076376	62.77	woody
Lauraceae	4	3000	0.003229127	119.26	woody
Liliaceae	15	4200	0.002767823	114.42	herb
Magnoliaceae	3	225	0.003522684	116.56	woody
Malvaceae	7	2300	0.001549981	33.9	woody
Oleaceae	6	600	0.001490366	59.5	herb/woody
Platanaceae	1	8	0.004305503	112.5	herb
Poaceae	35	9000	0.001761342	99.23	herb
Ranunculaceae	12	1700	0.003229127	130.435	woody
Rosaceae	57	4000	0.001684762	85.05	woody
Rubiaceae	6	13000	0.001614564	62.5	herb/woody
Rutaceae	1	1600	0.001614564	46	herb
Saxifragaceae	6	460	0.002039449	106.96	herb/woody
Solanaceae	7	2800	0.001549981	78.6	woody
Vitaceae	16	850	0.002583302	101.66	woody

Table 4 Species richness of Nearctic gall-inducing midges (Diptera: Cecidomyiidae) and associations with species richness, insularity, and habitat maximum height of plant genera within plant family Asteraceae.

Plant Genus	Species Richness Plant Genus	Plant Insularity	Plant Genus Maximum Height (cm)	Number of Cecidomyiid Species
Achillea	4	0.001484681	80	2
Ageratina	14	0.001608405	220	3
Ageratum	4	0.001378633	120	1
Ambrosia	22	0.001754623	90	5
Antennaria	34	0.001838177	50	3
Anthemis	2	0.001484681	90	1
Arctium	3	0.002270689	250	1
Artemisia	50	0.001484681	300	39
Bidens	25	0.00214454	400	3
Chondrilla	1	0.002412607	150	1
Cirsium	62	0.00214454	400	3
Conoclinium	3	0.001378633	200	1
Conyza	4	0.001378633	350	8
Coreopsis	28	0.002270689	50	1
Erigeron	173	0.001678335	100	3
Eupatorium	24	0.001331094	200	3
Flaveria	8	0.001754623	200	1
Grindelia	18	0.001429693	250	2
Helenium	18	0.00214454	160	3
Helianthus	52	0.001678335	380	9
Hymenopappus	10	0.001608405	120	2
Liatris	37	0.001331094	180	1
Palafoxia	10	0.001429693	60	1
Parthenium	7	0.002757265	120	1
Perityle	35	0.001544069	45	1
Pluchea	9	0.002031669	200	1
Rudbeckia	23	0.001608405	300	3
Senecio	55	0.002031669	250	3
Smallanthus	1	0.001838177	300	1
Solidago	77	0.001484681	200	22
Sonchus	5	0.002412607	200	1
Vernonia	17	0.002031669	20	2

Figure 1 Equilibrium predictions of island biogeographic theory to the relationship between insects and their host plants. (A) Colonization and extinction equilibria of speciose plant lineages relative to species poor plant lineages. (B) Colonization and extinction equilibria of genetically similar plants relative to genetically distinct plants. (C) Predicted equilibria between within plant speciation for genetically distinct plants and genetically similar plants, equilibrium points represent the proportion of surviving species resulting from within plant speciation.

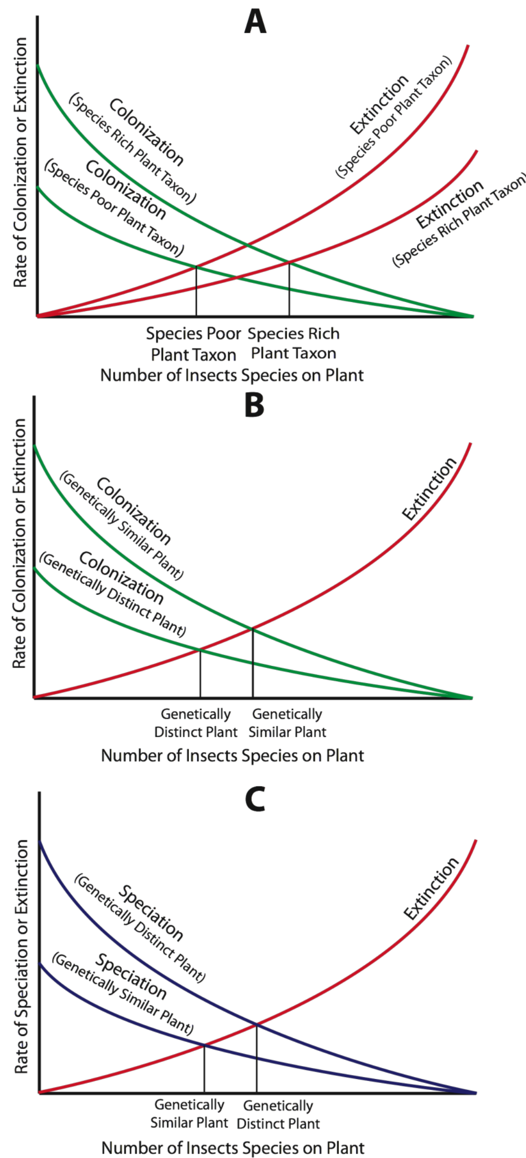


Figure 2 Predicted effect of plant insularity (genetic distinctiveness) on insect species richness. Increases in plant taxon (family, genus, or species) insularity (fewer close relatives) are predicted to be associated with decreases in the number of colonizing insects. Conversely, decreases in plant taxon insularity (many close relatives) are predicted to be associated with increases in insect diversity.

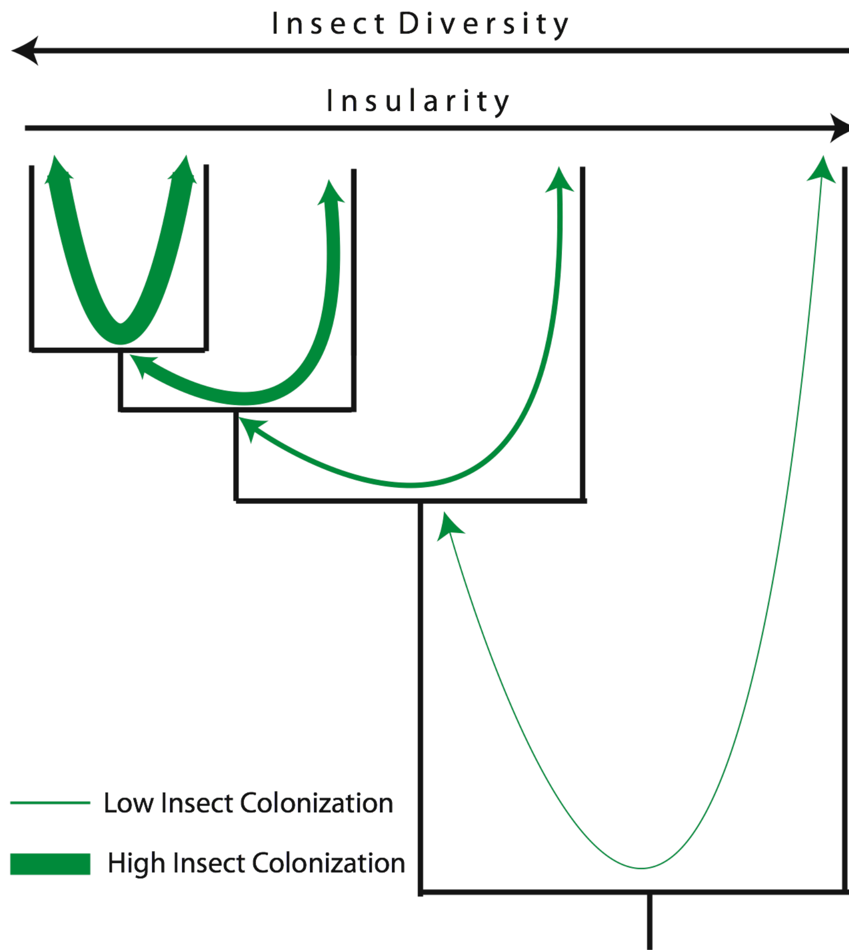


Figure 3 Idealized examples of predicted effects of insularity and ecological opportunity on diversity of phytophagous insects. (A) As insularity increases towards the right of the graph, diversity declines; diversity also declines towards the left of the graph with decreasing ecological opportunity. In this scenario, insect diversity peaks in the centre of the graph where ecological opportunity and insularity balance, illustrating the non-linear relationship that may result from interactions between insularity and ecological opportunity. (B) The effects of both decreasing insularity and increasing ecological opportunity, towards the left of the graph, result in maximal insect diversity; conversely as ecological opportunity declines and insularity increases towards the right of the graph insect diversity is minimized. Note that diversity is higher in (B) because ecological opportunity and insularity are operating in concert.

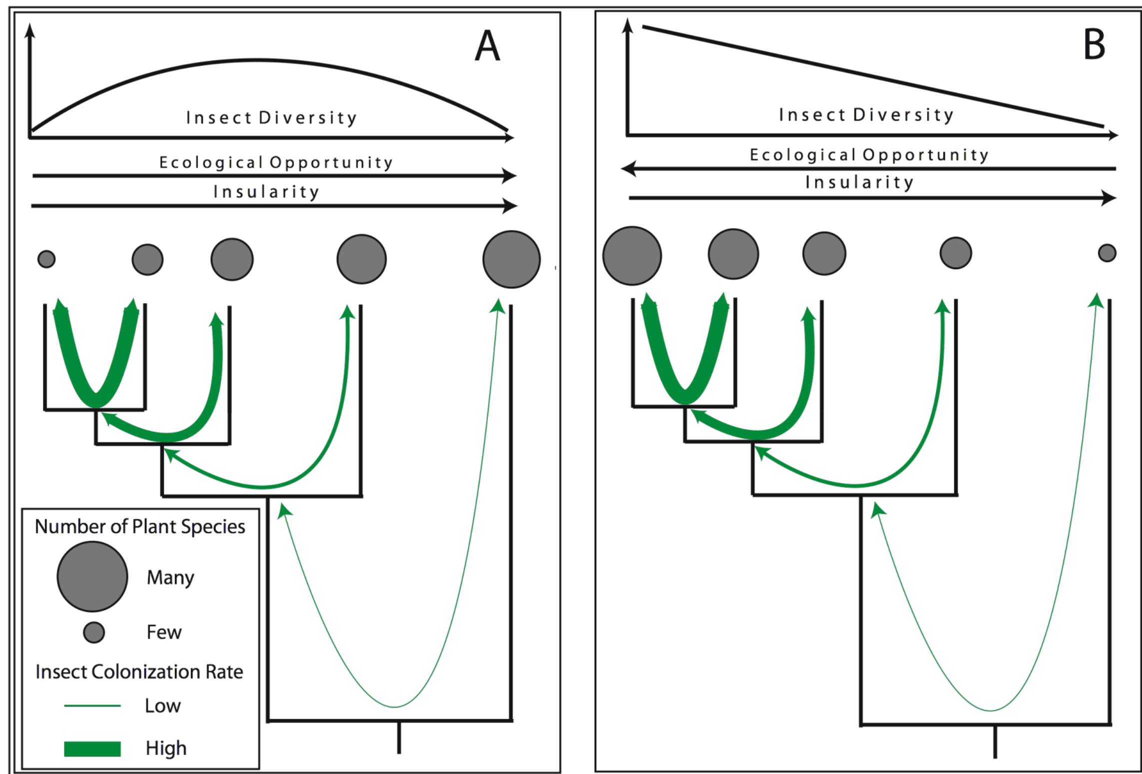


Figure 4 Relationship between plant insularity (x axis), plant species richness (y axis), and gall midge diversity (vertical axis): (A) decreases in plant family insularity ($\beta_{insularity}$= -2.759, p=0.006) and increases in plant family species richness ($\beta_{plant\ species\ richness}$=0.493, p=0.002) are associated with increases in Cecidomyiid species richness; (B) within Asteraceae decreases in plant genus insularity ($\beta_{insularity}$= -2.721, p=0.017) and increases in plant species richness ($\beta_{plant\ species\ richness}$= 0.584, p=0.001) are associated with increases in Cecidomyiid species richness.

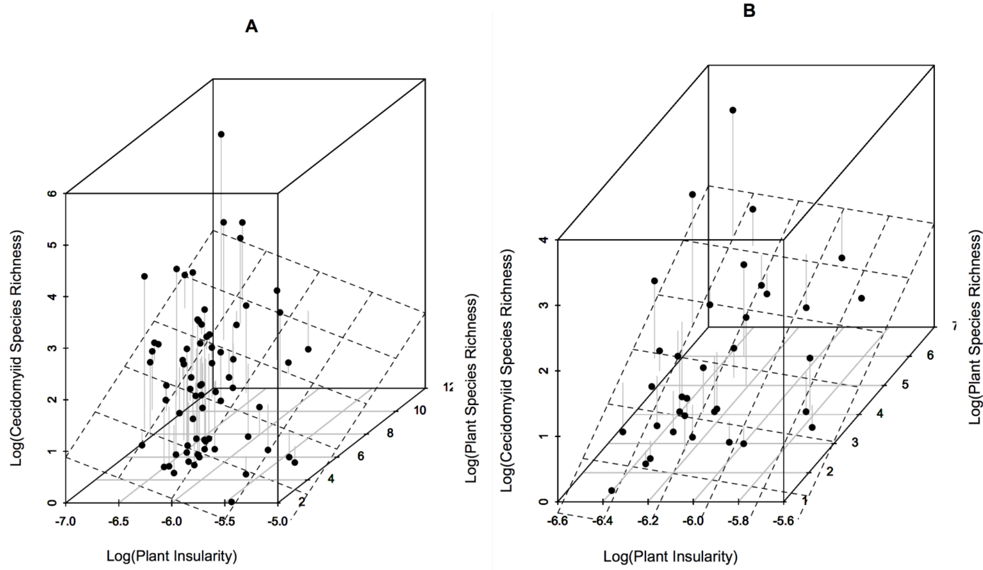


Figure 5 Relationship between plant species richness (y axis), plant family age (x axis) and gall midge diversity (vertical axis). Increases in plant family species richness ($\beta_{plant\ species\ richness}$=0.493, p=0.002) and plant family age (β_{age}= 3.832, p=0.002) are associated with an increase in gall midge species richness.

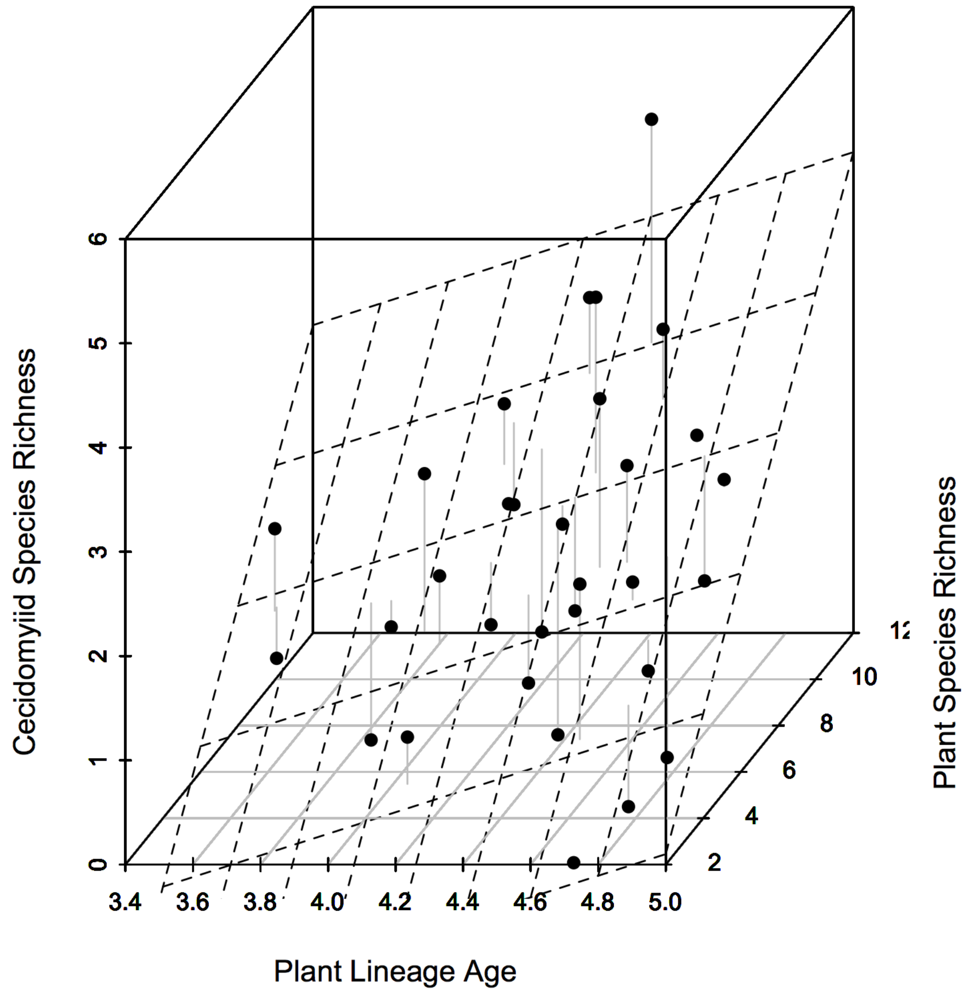
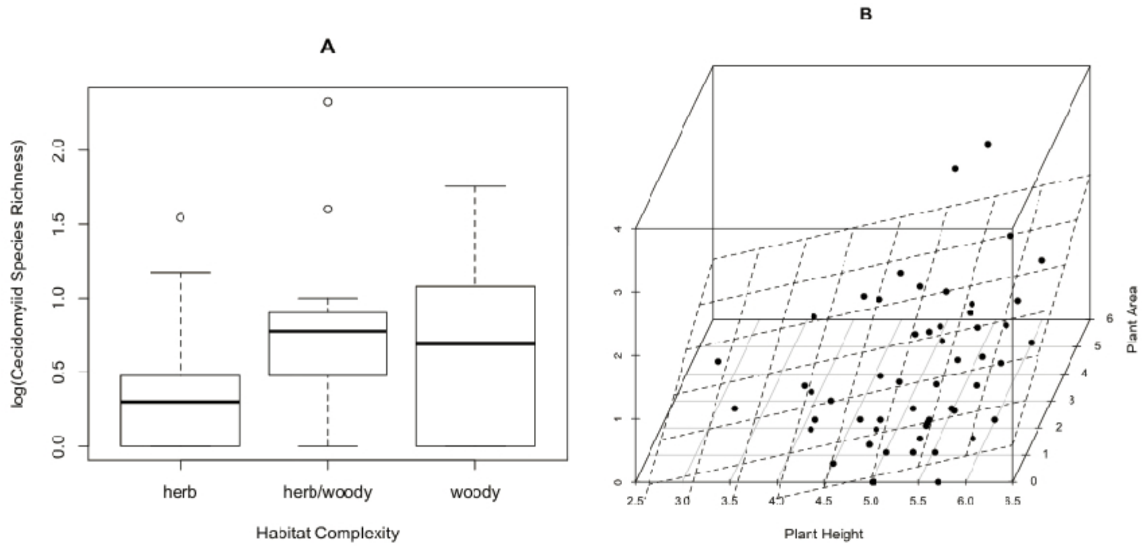


Figure 6 Relationship between gall midge diversity and plant habitat complexity: (A) host plant family habitat type (woody, herbaceous) vs. log gall midge species richness (R^2= 0.2554, $F_{5,37}$ = 3.881, p= 0.02; (B) Asteraceae host genera maximum height vs. log gall midge species richness (($\beta_{plant\ height}$= 0.811, p=0.016).



Chapter 3: Radiation of *Rhopalomyia* gall midges within the plant family Asteraceae

3.1 Abstract

Adaptive radiation of phytophagous insects is often associated with the invasion of new adaptive zones, defined as shifts to novel plant resources. Shifts may result in divergence in host-plant use, or divergence in the use of discrete plant resources within a single host-plant species. However, the evolutionary contexts in which shifts between, compared to shifts within, host-plant species are more important remain unclear. In this study I employ phylogenetic methods to characterize the radiation of *Rhopalomyia* gall midges in the context of a diverse, plant family (Asteraceae) with many close relatives. Analyses reveal that diversification of *Rhopalomyia* is associated with both shifts between host-plants and shifts to new niches within a host-plant species. As with other within host-plant radiations, speciation events appear to result from both divergence in plant-part use and divergence in life history timing. Such shifts to new plant resources likely exist along a continuum, with the probability of success defined by the opportunity for gene flow and the strength of natural selection required to colonize the new niche.

Keywords: adaptive radiation, within-host divergence, host shift, *Rhopalomyia*, phenology, gall morphology, ancestral state reconstruction, plant-part use.

3.2 Introduction

The underlying causes of diversification in plant-feeding insects are a subject of considerable interest in evolutionary biology (Condon et al. 2008; Janz et al. 2006). Speciation events in plant-feeding insects may be characterized as resulting from shifts between host-plant resources on a number of scales: between distantly related hosts, between closely related hosts, or between resources on the same host-plant species (Ehrlich and Raven 1964; Jermy 1984; Farrell and Mitter 1994; Dres and Mallet 2002; Nosil 2007; Després et al. 2002; Joy and Crespi 2007; Condon et al. 2008; Stireman et al. 2008). The relative importance of within-host versus between-host shifts may be a function of the host plant genetic distinctiveness, geographic range, and abundance within the range (Bernays and Chapman 1994; Crespi et al. 2004; Kelly and Southwood 1994). Thus, genetically distinct plants may experience fewer colonization events and relatively more within-plant speciation. Conversely, insects that live on host plant species with many close relatives may experience higher rates of colonization relative to within-plant speciation because open niches are often filled by colonization. Abundant, widely distributed host plant species may increase the frequency of encounter with insect species and thus may be larger colonization targets, thereby facilitating host shifts. Alternatively, widely distributed host-plant species may facilitate *in situ* speciation via greater chance that insect populations become isolated in a portion of a large host plant range.

Gall-inducing insects exhibit several characteristics which make them ideal for studying the processes driving the diversification of phytophagous insects. First, they are narrowly host specific, normally utilizing one plant species or a few closely related host-plant species (Shorthouse et al. 2005). Second, they are remarkably diverse, which

provides the comparative power necessary to draw inferences (Price 2005). Third, speciation events in gall-inducing insects are known to be associated with both shifts to new host-plant species (Price 2005) and *in situ* speciation (Joy and Crespi 2007; Stireman et al. 2008). In this study, I use a combination of mitochondrial and nuclear gene sequence data to evaluate hypotheses about the evolution of host-plant use in host-associated radiations of gall-inducing *Rhopalomyia* midges (Diptera: Cecidomyiidae) on two co-occurring, abundant and widespread host plant species in the plant family Asteraceae (big sagebrush (*Artemisia tridentata*) and rubber rabbitbrush (*Chrysothamnus nauseousus*)).

With more than 250 species *Rhopalomyia* is one of the largest known genera of gall midges occurring on every continent except Antarctica (Gagné 2004, Dorchin 2009). Most *Rhopalomyia* species are restricted to host plants within the plant family Asteraceae on which they induce galls on a great variety of plant parts (leaves, stems, nodes, buds, leaves, and flowers) and exploit different time periods of plant growth (Dorchin 2009).

Artemisia tridentata is the most abundant and widely distributed shrub in North America (McArthur et al. 1981). It is a large, woody, structurally complex evergreen shrub with silvery-grey leaves and yellow flowers (McArthur et al. 1981). The oldest North American fossil for *Artemisia tridentata* dates from the early Miocene, and molecular evidence supports the arrival of *Artemisia tridentata* in North America from Asia 10-18 million years ago (Tkach et al. 2008).

An evaluation of larval, pupal, adult, and gall characters of *Rhopalomyia* species associated with big sagebrush (*Artemisia tridentata*) has led to the identification of 26 species (Jones et al. 1983). These species all have very similar life histories, differing

mainly in their phenological relationships with the host-plant species (Jones et al. 1983). The phenotypic variation in galls induced by *Rhopalomyia* species is remarkable: galls may be discovered on all aerial plant growth including, leaves, stems, buds, leaf nodes, stem nodes, and flowers. Even within a plant part, galls of different species can be dramatically different in size, shape, colour, and pilosity. The ovipositing female deposits her eggs on the surface of the plant near, or on, the part to be infected, the gall is subsequently induced by the larva, and larval and pupal development occurs within the galls, with the adult stage lasting only hours (Jones et al. 1983).

Similarly to *Artemesia tridentata*, rabbitbrush (Asteraceae: *Chrysothamnus nauseousus*) is a perennial yellow-flowering shrub 30-230 cm tall, widely distributed throughout the arid regions of North America. Rabbitbrush is often associated with big sagebrush, and the plants are often found growing interdigitated. There are four distinct species of *Rhopalomyia* associated with rabbitbrush, each forming a unique gall on the stem or nodes of the plant (Gagné 1989).

Here I investigated the phylogenetic relationships among the groups of gall-inducing midges inhabiting big sagebrush and rubber rabbitbrush, to evaluate several hypotheses concerning their diversification. First, if diversification is associated with host shifts, then I expect the *Rhopalomyia* phylogeny to display a history of shifts among these and other related host plant species. Conversely, if within plant speciation has been the dominant mode in *Rhopalomyia* speciation, I expect that the *Rhopalomyia* phylogeny will form monophyletic groups associated with each host-plant species.

Second, among speciation events identified as occurring within host plant species, I hypothesize that, as with other within host-plant radiations (Cook et al. 2002; Joy and

Crespi 2007), divergence will be associated with plant-part shifts, phenological shifts, or both plant-part shifts and phenological shifts. To test inferences concerning divergence associated with shifts within a host plant, I considered the results of SH and Templeton tests (Shimodaira and Hasegawa 1999; Templeton 1983) comparing best trees with constraint trees that force taxa resulting from putative within host shifts to instead be sister to *Rhopalomyia* taxa on the nearest host-plant species.

To further evaluate contexts of gall midge divergence associated with shifts within a host plant, I evaluated the roles of changes in adult emergence phenology and divergence in plant-part use through sister taxa comparisons. If changes in adult emergence phenology are involved in mediating reproductive isolation between diverging gall midge taxa on the same host-plant species I expect sister taxa resulting from putative within-host speciation events to display divergence in adult emergence phenology. By contrast, if divergence in plant-part use has been an integral part of within host divergence process gall midge sister taxa on the same host-plant species are expected to show divergence in plant-part use. If both adult emergence phenology and divergence in plant part use were involved in the evolution of reproductive isolation between gall midge sister taxa, then sister taxa resulting from putative within host speciation events may show divergence in both adult emergence phenology and plant-part use. Alternatively, if neither divergence in plant part use or shifts in adult emergence timing were important during divergence, then sister taxa may not be diverged in either trait.

3.3 Methods

3.3.1 Collection Sites

Galls of *Rhopalomyia* species associated with big sagebrush (*Artemisia tridentata*) and with rabbitbrush (*Chrysothamnus nauseousus*) were collected from sites throughout the great basin desert in the south Okanagan of British Columbia Canada, and in Idaho and Washington states in the United States. Galls were stored individually in vials and transported to the laboratory where they were kept at room temperature until adults emerged. Upon emergence adults were preserved whole in 20% dimethyl sulphoxide in a saturated solution of NaCl and/or stored at -20 degrees Celsius. Additional *Rhopalomyia* specimens were obtained from the Smithsonian Institution National Museum of Natural History in Washington, DC.

3.3.2 Collection of DNA Data

Whole genomic DNA was isolated using standard phenol chloroform techniques (Hillis et al. 1996) from single adult midges of both sexes. DNA was extracted from as many individuals of each species as possible (Table 5). I amplified a 450 base pair fragment of the mitochondrial gene cytochrome oxidase subunit I (COI) using primers C1-J-1718 and C1-N-2191 (Simon et al. 1994). I also amplified a 550 base pair fragment of the wingless gene (Wg) using primers 5'wg1 and 3'wg2 (Ober 2003). Resulting PCR products were purified using the Qiaquick PCR Purification Kit (Qiagen Inc.). Purified PCR products were sequenced directly using Eurofins MWG Operon sequencing service. I obtained outgroup sequences for all other available *Rhopalomyia* species available on Genbank (Table 6).

3.3.3 Phylogenetic Analyses

Sequences were aligned using Clustal (Thompson et al. 1994) and adjusted by eye using Se-Al (Rambaut 1996). The best model of molecular evolution for each gene was estimated using both Modeltest (Posada and Crandall 1998) and MrModeltest (Nylander 2004) for use in maximum likelihood (ML) and Bayesian phylogeny reconstruction respectively. I inferred phylogenies for each gene separately using ML and Bayesian methods as implemented in PAUP* 4.0b10 (Swofford 2002) and MrBayes 3.12 (Ronquist and Huelsenbeck 2003). Support for inferred phylogenies was determined using 500 ML bootstrap replicates and Bayesian posterior probability values. To determine whether it was appropriate to combine data, I assessed heterogeneity of gene sequence data using the partition homogeneity test as implemented in PAUP* 4.0b10 (Swofford 2002). The combined dataset was analyzed in a 2 partition analysis applying the best fit model of molecular evolution to each partition in MrBayes 3.12 (Ronquist and Huelsenbeck 2003). Sequences from the most closely related genus (*Mayetiola destructor*) in the same sub-family (Cecidomyiinae) for which data were available on Genbank was used as an outgroup taxon.

Inferences about the evolution of host plant relationships among *Rhopalomyia* species are complicated by the large size of the genus, which comprises 250 species world wide, 90 in North America, and 86 on plants in family Asteraceae. This study includes data from 23 North American *Rhopalomyia* species and 5 species from Korea, Japan, and China. I evaluated monophyly of clades of *Rhopalomyia* forming galls on sagebrush and rabbitbrush through nodal support values (ML and Bayesian). I also considered the results of Templeton and Shimodaira-Hasegawa tests (SH) tests

implemented in PAUP* 4.0b10 (Templeton 1983; Shimodaira and Hasegawa 1999; Swofford 2002) comparing the best trees to constraint trees which forced the monophyly of each host-plant associated group of *Rhopalomyia*. To test inferences of within plant speciation, I considered the results of nodal support values for sister pairs inferred to be derived from within host-plant speciation events. To further evaluate hypotheses concerning the roles of *Rhopalomyia* adult emergence phenology and plant-part shifts in divergence of species pairs putatively resulting from within host plant speciation events, I performed sister taxa comparisons.

3.4 Results

3.4.1 Dataset

The dataset of cytochrome oxidase subunit I (COI) and wingless (Wg) nucleotide sequences for 28 *Rhopalomyia* species consisted of 1000 positions (450 COI, 550 Wg). The region of Wg amplified contains a small (126 base pairs) intron. Of the 1000 positions, 468 were parsimony informative (207 COI, 167 Wg). Interspecific pairwise differences within the ingroup ranged from 2.14% to 24.81% for COI and 1.10% to 20.08% for Wg. Differences between the ingroup and outgroups were 14.57% to 29.9% for COI and 10.06% to 23.11% for Wg. A partition homogeneity test showed that the two gene regions were not incompatible ($P > 0.05$).

3.4.2 Phylogentic Inferences and Evolution of Host-plant Use

The phylogenies for the two gene regions showed only minor differences (Figure 7), the main difference being placement of *R. chrysothamnus* within the *Artemisia* feeding *Rhopalomyia* in the topology derived from the *wingless* gene. The more rapidly-

evolving *COI* locus shows that each species formed well-supported monophyletic groups, the sole exception being *Rhopalomyia pomum* which is paraphyletic with respect to *Rhopalomyia calvipomum.* The phylogeny (Figure 8) also shows several other notable patterns. First, most of the sampled *Rhopalomyia* midges inducing galls on big sagebrush form a well-supported clade (posterior probability 0.92) relative to other North American and Asian *Rhopalomyia* species, but with *Rhopalomyia chrysanthemumi* from North American *Chrysanthemum* included with the *Artemisia* feeding species. Second, three other *Rhopalomyia* species from sagebrush (*R. florella, R. lignitubus,* and *R. obovata*) appear more closely related to *Rhopalomyia* species from other host plants (e.g. *Chrysothamnus*). Phylogenies constraining all *Rhopalomyia* which induce galls on big sagebrush to be monophyletic were significantly worse than best trees under MP using Templeton's test (difference in length = 31, $P < 0.001$) and under ML as judged by Shimodaira-Hasegawa's tests (difference in –In L = 61.85, P = 0.028).

The three sampled species inducing galls on rubber rabbitbrush form a well supported clade (posterior probability 1.0) with *Rhopalomyia* which induce galls on North American late goldenrod (*Solidago altissima*) and grass leaved goldenrod (*Euthamia graminifolia*). Constraint trees which forced the *Chrysothamnus Rhopalomyia* to be monophyletic were not significantly worse than best trees under MP using Templeton's test (difference in length = 6, P = 0.56) or ML using Shimodaira-Hasegawa's test (difference in –In L = 13.92, P = 0.22).

Taken together the most likely phylogenies, ML and Bayesian nodal support values, and results of SH and Templeton topology tests do not support monophyletic groups of *Rhopalomyia* confined to each host-plant species but do support pairs of sister

taxa originating through *in situ* speciation on the same host plant, with respect to the other sampled *Rhopalomyia* taxa.

3.4.4 Ecological Contexts of Within-host Speciation

3.4.4.1 Evolution of Plant Part Use

Bud galling is the most common state among the sampled species (17 of 28), and this form of gall-induction is also widely distributed between clades (Figure 8). Thus one obvious hypothesis derived from the phylogeny concerning the evolution of plant-part use entails evolution of bud galling in the ancestor of the group, followed by gains of galling nodes, leaves, and stems (Figure 8). Three other notable inferences concerning the evolution of plant-part use can be made from the phylogeny. First, the most likely phylogeny and results of Templeton (difference in length = 10, $P = 0.04$) tests suggest that node galling evolved multiple times among the sampled species (*R. medusa* and *R. nucula*) from bud galling ancestors. However, SH tests (difference in –In L = 22.75, P = 0.06) showed that constraint trees forcing node galling species together were (marginally) not significantly worse than the most likely phylogeny. Second, species that induce galls on the leaves of *Artemisia tridentata* predominantly form a well-supported monophyletic group (posterior probability of 1.0) suggesting that within host-speciation may be important among leaf galling taxa. One other leaf galling *Rhopalomyia* species (*Rhopalomyia brevibulla*) groups with two bud-galling species (*R. cramboides* and *R. conica*) suggesting at least 2 origins of leaf galling.

3.4.4.2 Evolution of Life History Timing

Rhopalomyia may be uni-voltine, bi-voltine, or multi-voltine (Dorchin 2009), and most species found on big sagebrush are uni-voltine (Jones et al. 1983). Extensive collections performed by Jones et al. (1983) showed that *Rhoplaomyia* inducing galls on big sagebrush have very similar life histories, differing mainly in their phenological relationships with the host plant. Sister taxa comparisons among uni-voltine species galling the same part (bud) of *Artemisia tridentata* show that two pairs of sister taxa apparently diverged in life history timing, *R. anthoides* and *R. lignea*, and *R. florella* and *R. lignitubus* (Figure 8). Each pair is diverged in adult emergence timing (approximately 1 month) such that adults which live only hours (Jones et al. 1983) are unlikely to encounter adults of their sister taxon.

3.5 Discussion

Phylogenetic analyses of multiple genes reveal that host-plant use among *Rhopalomyia* species has evolved through a combination of switching between related host-plant species (for example, the apparent shift from *Artemisia tridentata* to *Chrysanthemum spp*.) and within plant speciation. Further, SH and Templeton tests fail to support the monophyly of host-plant use hypothesis for *Rhopalomyia* on big sagebrush, consistent with the hypothesis that diversification of *Rhopalomyia* among plants in family Asteraceae involves some degree of host plant switching.

A key limitation of this study is the incomplete sampling of the *Rhopalomyia* species on *Artemisia* and related host-plant species. Appropriate and extensive taxon sampling is one of the most important determinants of accurate phylogeny estimation (Heath et al. 2008; Hillis et al. 2003). Thus, inference and interpretation drawn from a

phylogeny are also strongly biased by incomplete taxon sampling. In this study, inferences concerning both the evolution of host-plant preferences and processes which may be involved in putative within host speciation events would be strengthened by the addition of more *Rhopalomyia* species from *Artemisia tridentata* and other host-plant species. Specifically, inferences regarding the involvement of life-history timing and divergence in plant-part use in within host speciation are heavily predicated on the state of their sister taxon. The addition of further species, with differing or the same state, may break up associations between taxa currently inferred to be sister altering inferences about the role of life history timing or plant-part use. Additionally, more taxa from other host plant species may further divide the two main clades of *Rhopalomyia* feeding on *Artemisia* thus increasing the number of inferred host-plant shifts.

All of the *Rhopalomyia* used in this study were collected from related plant genera, or species, within a single large plant family (Asteraceae). Host-plant shifts require adaptation of the colonizing insect species to the nutritional chemistry, plant defences, and phenology of the novel environment, as well as to potentially new natural enemies (Farrell and Mitter 1994; Cooke et al. 2002; Jaenike 1989; Nosil 2007). However, the probability of success of such shifts between plant resources may be continuous. The probability of divergence is thus likely a balance between the strength of natural selection, highest among distantly related host plant species, and opportunity for gene flow, highest during shifts to resources within the same host plant species (Slatkin 1987; Lenormand 2002; Nosil 2002; Joy and Crespi 2007). Shifts among closely related plant species in speciose plant families may therefore provide optimum conditions for divergence, whereby selection is strong but the homogenizing effects of gene flow

may be reduced. By this hypothesis, I expect the rate of adaptive radiation associated with shifts to new niches within the same host plant to proceed more slowly than adaptive radiation that is predominantly associated with shifting among related host-plant species. Shifts among resources in the same host plant likely require substantially less adaptation because many of the selective forces will be relatively similar to the originating population. Thus, selection against hybrids may be reduced between populations on the same host plant species because the strength of divergent selection is expected to be lower and the opportunity for gene flow is expected to be higher within host relative to between host. This idea could be tested by comparing speciation rates among radiations of phytophagous insects which have speciated predominantly through host-plant shifts with those which have speciated through within host-plant speciation.

Among the North American *Rhopalomyia* feeding upon *Artemisia tridentata,* within host speciation events appear to be derived from a combination of divergence in plant-part use (shifts between plant organs, i.e. bud to node) and divergence in life history timing (Figure 8; Jones et al. 1983). The processes of divergence in life history timing and plant-part use have been shown to be important in within host plant adaptive radiations of *Andricus* gall wasps (Cook et al. 2002), *Blephanoneura* fruit flies (Condon et al. 2008), *Asphondylia* gall midges (Joy and Crespi 2007), *Strobilomyia* cone flies (Sachet et al. 2006), and *Chiastochaeta* globe flower flies (Després et al. 2002). In sympatry, these two processes likely promote the evolution of reproductive isolation rapidly when operating together because, factoring out temporal variation in natural enemies, divergent selection is likely to be relatively higher between parts of a host plant than between time periods on the same plant part. Furthermore, the opportunities for

gene flow are likely to be reduced between time periods relative to between plant parts at the same time period.

Taken together, the results of this study provide evidence that adaptive radiation in *Rhopalomyia* gall midges proceeds through a combination of host-plant shifts and within host-plant speciation. The sequencing of more *Rhopalomyia* species feeding upon *Artemisia* and related host plants would allow further testing of hypotheses regarding the evolution of host plant use among *Rhopalomyia* species and more clearly elucidate patterns associated with within host-plant speciation.

3.6 Acknowledgements

This manuscript was greatly improved by comments of Felix Breden and Arne O. Mooers. I thank Tanja R. Schwander, Thierry Heger, Timothy W. Joy, Forrest Joy, and the SFU FAB* lab for assistance with collecting field samples. I am grateful to Dr. Ray Gagné of the Smithsonian National Museum for providing *Rhopalomyia* specimens in alcohol. I thank the Canadian Natural Sciences and Engineering Research Council of Canada (NSERC), and the Society for Systematic Biology for financial support.

3.7 Literature Cited

Bernays, E.A., and Chapman, R.F. 1994. *Host plant selection by phytophagous insects.* Chapman and Hall.

Condon, M.A., S.J. Scheffer, M.L. Lewis, and S.M. Swensen. 2008. Hidden neotropical diversity: greater than the sum of its parts. *Science* 320:928-931.

Cook, R. J., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gallwasps. Evolution 56:1821-1830.

Crespi, B. J., D.C. Morris, and L.A. Mound. 2004. *Evolution of Ecological and Behavioural Diversity: Australian Acacia Thrips as Model Organisms*. Australian Biological Resources Study & CSIRO Entomology: Canberra, Australia, 328 pages.Dorchin, N. 2009. Revision of the goldenrod-galling *Rhopalomyia* species (Diptera: Cecidomyiidae) in North America. Zootaxa 2152: 1-35.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. (2002) Speciation in the globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258–268.

Dres, M. and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. B 357: 471-492.

Ehrlich, P.R. and P.H. Raven. 1964. Butterflies and plants: a study in co-evolution. Evolution 18:586-608.

Farrell, B.D. and C. Mitter. 1994. Adaptive radiation in insects and plants: time and opportunity. Amer. Zool. 34:57-69.

Gagné R.J. 1989. The plant-feeding gall midges of North America. Cornell University Press, Ithaca.

Heath, T.A., S.M. Hedtke, D.M. Hillis. 2008. Taxon sampling and the accuracy of phylogenetic analyses. Journal of Systematics and Evolution 46: 239–257.

Hillis, D.M., D.D. Pollock, J.A. McGuire, and D.J. Zwickl. 2003. Is sparse taxon sampling a problem for phylogenetic inference? Systematic Biology 52:124-126.

Hillis, D. M., C. Moritz, and B. K. Mable. 1996. Molecular Systematics (2nd Edition). Sinauer and Associates.

Jaenike, J. 1989. Genetics of butterfly host-plant associations. Trends. Ecol. Evol. 4:34-35.

Janz, N., S. Nylin, and N. Wahlberg. 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. BMC Evol. Biol. 6: 4.

Jermy,T. 1984. Evolution of insect/host plant relationships. Am. Nat. 124:609-630.

Jones, R. G., R. J. Gagné, and W. F. Barr. 1983. Biology and taxonomy of the *Rhopalomyia* gall midges (Diptera: Cecidomyiidae) of *Artemesia tridentata* Nutall (Compositae) in Idaho. Contrib. Am. Entomol. Instit. 21:1-76.

Joy J.B., B.J. Crespi. 2007. Adaptive radiation of gall-inducing insects within a single host-plant species. Evolution 61:784-795.

Kelly, C.K. and T.R.E. Southwood. 1999. Species richness and resource availability: A phylogenetic analysis of insects associated with trees. PNAS 96: 8013–8016.

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17:183-189.

McArthur, E.D., C.L. Pope, and D.C. Freeman. 1981. Chromosomal studies of subgenus Tridentatae of *Artemisia*: evidence for autopolyploidy. Am. J. Bot. 68: 589-605.

Nosil, P., B.J. Crespi, and C.P. Sandoval. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417:440-443.

Nosil, P. 2007. Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. Am. Nat. 169:151-162.

Nylander JA. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala Univ, Sweden.

Ober, K. 2003. Phylogenetic relationships of the carabid subfamily Harpalinae (Coleoptera) based on molecular sequence data. Mol. Phylogenet. Evol. 26:334-336.

Posada D, Crandall KA. 1998. MODELEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Price, P.W. 2005. Adaptive radiation of gall-inducing insects. Basic and Applied Ecology 6:413-421.

Rambaut A. 1996. Se-Al: Sequence alignment editor. Available at http://evolve.zoo.ox.ac.uk/.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Shimodaira, H. and Hasegawa, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **16**, 1114–1116.

Shorthouse, J.D., D. Wool, and A. Raman. 2005. Gall-inducing insects – nature's most sophisticated herbivore. Basic and Applied Ecology 6:407-411.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook 1994, Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a

compilation of conserved polymerase chain reaction primers. Annu. Rev. Entomol. 87: 651-701.

Slatkin, M. 1987. Gene flow and geographic structure of natural populations. Science 236:787-792.

Stireman J.O, E.M. Janson, T.G. Carr, H. Devlin, and P. Abbot. 2008. Evolutionary radiation of *Asteromyia carbonifera* (Diptera: Cecidomyiidae) gall morphotypes on the goldenrod *Solidago altissima* (Asteraceae). Biol J Linn Soc Lond. 95:840-858.

Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony(*and other methods). Ver. 4. Sinauer Associates, Sunderland, MA.

Templeton, A. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221-224.

Thompson J.D, D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc Acids Res 22:4673-4680.

Tkach, N.V., M.H. Hoffmann, M. Roser, and K.B. von Hagen. 2008. Temporal patterns of evolution in the Arctic explored in *Artemisia* L. (Asteraceae) lineages of different ages. Plant Ecology and Diversity 1:161-169.

Table 5 Details of sequence data for each species for each gene used in this study and sampling locations.

Species	COI	Wg	Host Plant	Plant Part	Sampling Location
R. pomum	2	1	*Artemisia tridentata*	Leaf	North America
R. calvipomum	2	2	*Artemisia tridentata*	Leaf	North America
R. hirtipomum	4	4	*Artemisia tridentata*	Leaf	North America
R. hirtibulla	1	1	*Artemisia tridentata*	Leaf	North America
R. brevibulla	2	2	*Artemisia tridentata*	Leaf	North America
R. ampullaria	1	0	*Artemisia tridentata*	Leaf	North America
R. lignea	4	3	*Artemisia tridentata*	Bud	North America
R. lignitubus	4	2	*Artemisia tridentata*	Bud	North America
R. cramboides	4	4	*Artemisia tridentata*	Bud	North America
R. anthoides	4	2	*Artemisia tridentata*	Bud	North America
R. florella	4	4	*Artemisia tridentata*	Bud	North America
R. conica	4	0	*Artemisia tridentata*	Bud	North America
R. obovata	1	1	*Artemisia tridentata*	Bud	North America
R. nucula	2	1	*Artemisia tridentata*	Node	North America
R. medusa	5	3	*Artemisia tridentata*	Node	North America
R. chrysothamni	1	1	*Chrysothamnus nauseosus*	Stem	North America
R. utahensis	2	2	*Chrysothamnus nauseosus*	Bud	North America
R. sp.	1	1	*Chrysothamnus nauseosus*	Bud	North America
R.chrysanthemi	1	0	*Chrysanthemum*	Flower	North America
R. solidaginous	1	0	*Solidago altissima*	Bud	North America
R. lobata	1	0	*Euthamia graminifolia*	Bud	North America
R. fusiformae	1	0	*Solidago juncea*	Leaf	North America
R. shinjii	1	0	*Artemisia montana*	Bud	Japan/Korea
R. yomogicola	1	0	*Artemisia*	Bud	Korea/Japan
R. foliorum	1	0	*Artemisia princeps*	Leaf	Japan
R. struma	1	0	*Artemisia*	Bud	Korea/Japan
R. longicauda	1	0	*Chrysanthemum*	Bud	Korea/Japan/China
R. sp.	1	0	*Artemisia*	Bud	Korea/Japan

Table 6 Accession numbers for *Rhopalomyia* sequences obtained from Genbank.

Species	**Accession**
R.lobata	gi82466892
R.solidaginous	gi82466954
R.fusiformae	gi82466894
R.longicauda	gi224176018
R.yomogicola	gi163929870
R.foliorum	gi157144123
R.shinjii	gi151175659
R.struma	gi46020127
R.sp	gi46020067
Mayetiola	EU375697

Table 7 *Rhopalomyia* gall morphological characteristics and adult emergence phenology on *Artemisia tridentata.*

Species	**Adult Emergence Phenology**	**Mean Gall Length (mm)**	**Mean Gall Width (mm)**	**Mean No. Larval Capsules**	**Max No. Larval Capsules**
R. pomum	April 23-June13	26	20	3.7	17
R. calvipomum	April 19-28	15	13	>3	>5
R. hirtipomum	May 14-July 11	12	11	1	1
R. hirtibulla	March 27-29	1	1	1	1
R. brevibulla	March 23-25	1	1	1	1
R. ampullaria	July 20-August 5	1	2	1	1
R. lignea	April 22-May 15	4.7	7.3	>2	>5
R. lignitubus	May 14-19	3	19	1	1
R. cramboides	April 22-25	9	8	3	8
R. anthoides	March 21-29	8	11	4	49
R. florella	March 25-26	-	-	-	>1
R. conica	April 20-May16	4	3.2	1	1
R. obobata	January, March, May	7	8	-	4
R. nucula	March 20-April 19	9	6	5	11
R. medusa	April 20-July 11	13.3	13.3	1.2	3

Figure 7 *Rhopalomyia* phylogenies for A. Cytochrome oxidase subunit I (COI) gene including all North American and Asian species for which data are available. B. Wingless (WG) gene for samples collected within North America.

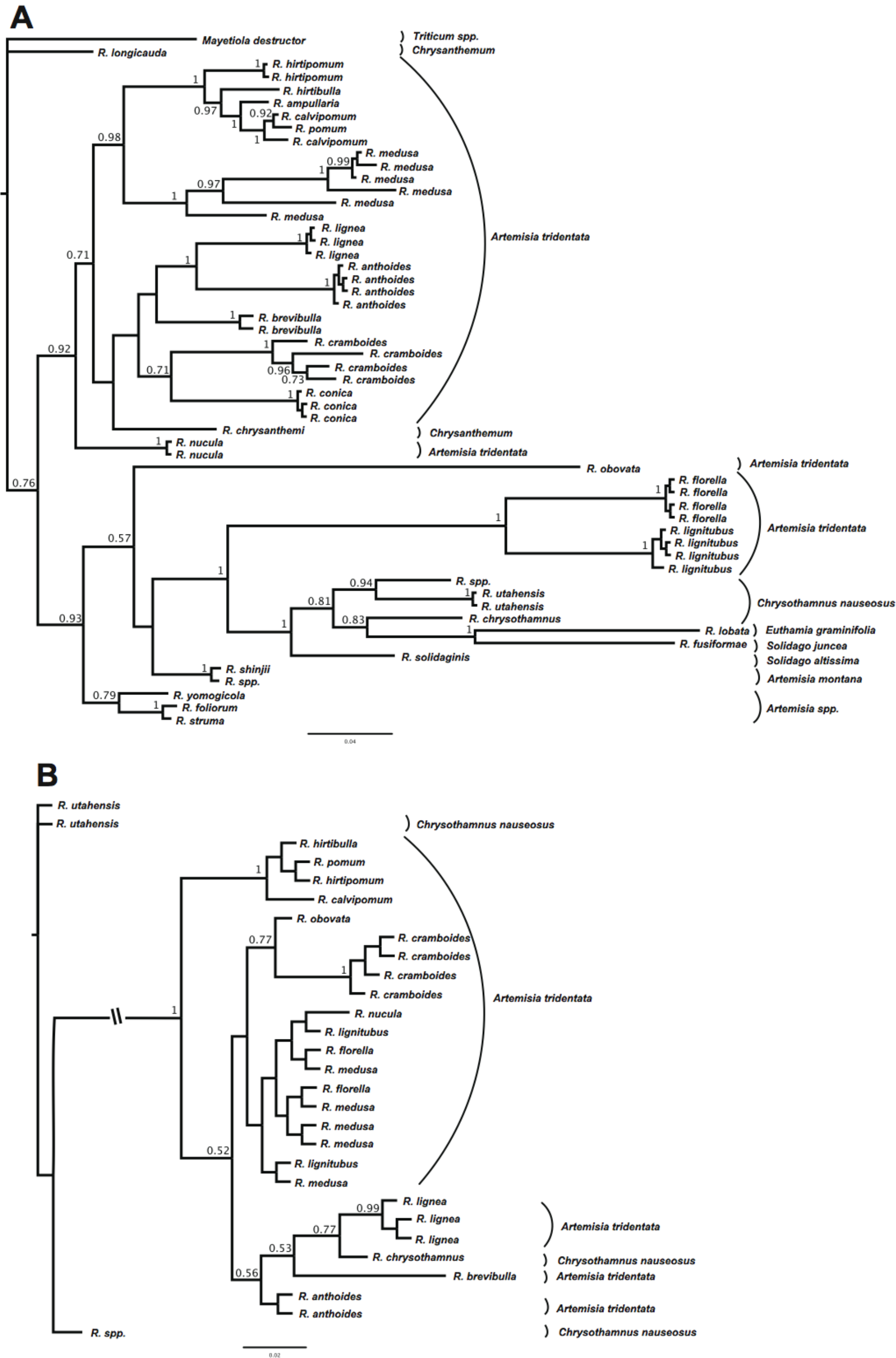
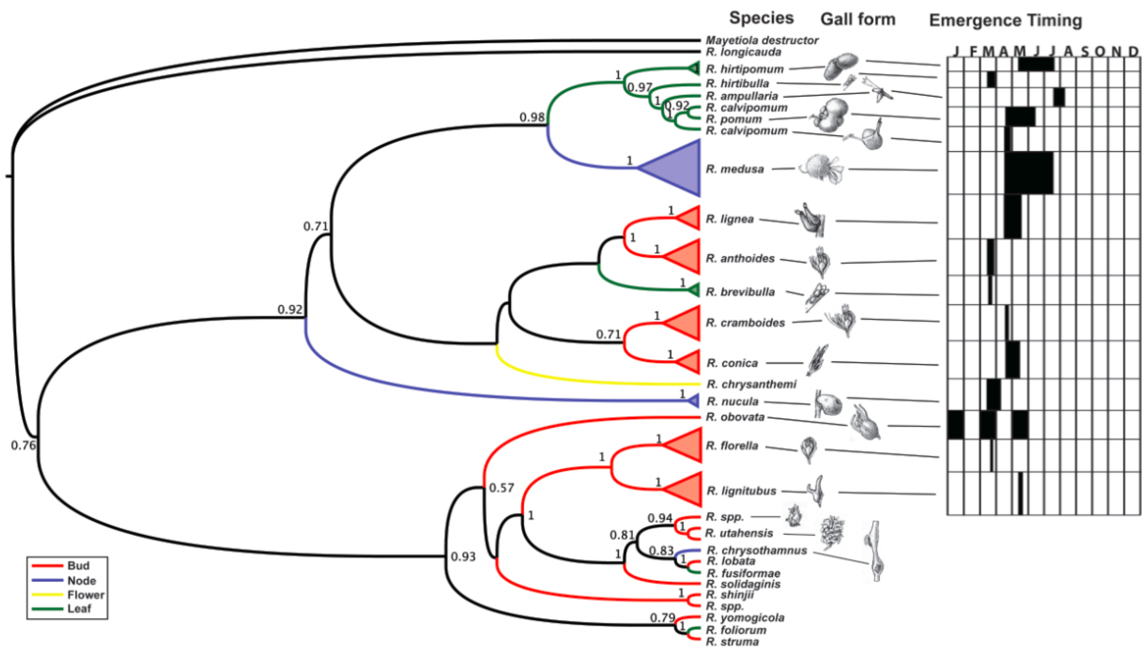


Figure 8 Combined COI and WG phylogeny for *Rhopalomyia* gall midge species. Life history data (phenology, gall morphology) for each species on big sagebrush (*Artemisia tridentata*) is presented at the tips.



Chapter 4: Adaptive radiation of gall-inducing insects within a single host plant species

Modified from Joy and Crespi (2007) by permission of Blackwell Publishing.

4.1 Abstract

Speciation of plant-feeding insects is typically associated with host-plant shifts, with subsequent divergent selection and adaptation to the ecological conditions associated with the new plant. However, a few insect groups have apparently undergone speciation while remaining on the same host plant species, and such radiations may provide novel insights into the causes of adaptive radiation. We used mitochondrial and nuclear DNA to infer a phylogeny for 14 species of gall-inducing *Asphondylia* flies (Diptera: Cecidomyiidae) found on *Larrea tridentata* (creosote bush), which have been considered to be monophyletic based on morphological evidence. Our phylogenetic analyses provide strong support for extensive within-host plant speciation in this group, and it demonstrates that diversification has involved numerous shifts between different plant organs (leaves, buds, flowers, and stems) of the same host-plant species. Within-plant speciation of *Asphondylia* is thus apparently facilitated by the opportunity to partition the plant ecologically. One clade exhibits temporal isolation among species, which may have facilitated divergence via allochronic shifts. Using a novel method based on Bayesian reconstruction, we show that the rate of change in an ecomorphological trait, ovipositor length, was significantly higher along branches with inferred shifts between host plant organs than along branches without such shifts. This

finding suggests that *Larrea* gall midges exhibit close morphological adaptation to specific host plant parts, which may mediate ecological transitions via disruptive selection.

Keywords: Adaptive radiation, *Asphondylia*, ecological shifts, galling, insect-plant interactions, plant-part specific specialization, speciation.

4.2 Introduction

Plant-feeding insects have several characteristics that make them useful models for the study of speciation. First, the high diversity of phytophagous insects and the continuum of populations exhibiting various stages of reproductive isolation facilitate comparative analyses of speciation mechanisms (Drès and Mallet 2002). Second, most phytophagous insects are ecologically specialized on particular host-plant resources, and such specialization may facilitate the evolution of reproductive isolation (Jaenike 1989; Caillaud and Via 2000). Third, the developmental timing of phytophagous insect populations can be determined by host-plant resources with different phenologies, such that adults from populations specialized on different host-plant resources may mature and mate at different times, leading to temporal isolation (Feder and Filchak 1999; Groman and Pellmyr 2000).

Shifts to new host plant species have played a crucial role in the diversification of phytophagous insects (Ehrlich and Raven 1964; Jermy 1984; Farrell and Mitter 1994; Thompson 1994; Mardulyn, Milinkovitch et al. 1997; Becerra and Venable 1999; Funk, Filchak et al. 2002). Speciation via host shifting often proceeds via the development of prezygotic isolation, associated with fidelity of mating on the host plant (Berlocher 2000;

Feder, Berlocher et al. 2003). Such prezygotic isolation can lead to the formation of host plant races exhibiting moderate levels of reproductive isolation, and in time these host races may differentiate into species (Drès and Mallet 2002). Such host-plant shifts and the evolution of host races have been proposed as a common scenario for non-allopatric speciation (Craig, Itami et al. 1993; Feder, Opp et al. 1994; Futuyma, Keese et al. 1995; Berlocher 2000; Groman and Pellmyr 2000; Abrahamson, Eubanks et al. 2001; Craig, Horner et al. 2001; Emelianov, Dres et al. 2001; Dres and Mallet 2002), although strong support for these mechanisms has remained elusive.

Recent phylogenetic and ecological studies of several clades of phytophagous insects have demonstrated that speciation can also occur in the absence of host plant shifts (Condon and Steck 1997; Cook, Rokas et al. 2002; Després, Pettex et al. 2002). In these cases, speciation is often associated with shifting to different parts of the same host plant species, such as from leaf to stem, and the evolution of reproductive isolation may often involve phenological separation (Condon and Steck 1997; Després, Pettex et al. 2002; Ferdy, Després et al. 2002). These patterns of within-host speciation are also not limited to phytophagous insects: for example, Simkova et al. (2004) showed that in a group of monogean parasites of fishes, diversification is explained in part by within-host speciation. Cases of within-host speciation may provide useful insights into speciation, because in these cases the effects of ecology on divergence are likely easier to partition from alternative processes, and divergence may be more likely to involve non-allopatric processes in the evolution of reproductive isolation.

Gall midges (Diptera: Cecidomyiidae) are unusual among phytophagous insects in that taxonomic classifications show that many genera exhibit large groups of putatively

closely-related species found on a single host-plant species (Jones, Gagne et al. 1983; Hawkins, Goeden et al. 1986; Gagne 1989; Gagne and Waring 1990). Gall midges comprise the largest radiation of galling insects (Ronquist and Liljeblad 2001). They form galls on virtually all plant parts (leaves, stems, twigs, buds, flowers, and roots). Cecidomyiids are widely distributed among host plants, occurring on gymnosperms, angiosperms, monocotyledons and dicotyledons (Gagne 1989). Most cecidomyiids, like other gall-inducing insects (Crespi, Carmean et al. 1997) are highly host-plant specific, most often feeding only on one part of a single host plant species (Jones, Gagne et al. 1983; Hawkins, Goeden et al. 1986; Gagne 1989). For example, within the large genus *Asphondylia* (247 described species world wide), members of morphologically-based species groups, defined by similarities in larval, pupal, and adult characters, are often associated with the same host plant species (Hawkins, Goeden et al. 1986; Gagne and Waring 1990).

Current understanding of phylogenetic relationships amongst the Cecidomyiidae is highly incomplete, such that patterns of host associated radiations in this group remain largely unexplored (Dorchin et al. 2004). Based on larval, pupal, and adult morphological characters, gall-inducing flies of the *Asphondylia auripila* group are believed to form a monophyletic group in which all of the species feed upon creosote bush (*Larrea tridentata*) (Waring and Price 1989). Creosote bush displays a disjunct distribution covering vast areas of the North American and South American mid-latitude arid regions. The *Asphondylia auriplila* group is thought to have arisen following the arrival of creosote bush in North America as no *Asphondylia* species where found inhabiting South American creosote bush (Gagne and Boldt 1991). Members of the *Asphondylia auripila*

group differ in several ecologically important characteristics such as gall morphology, gall position, and ovipositor characteristics. The life histories of these midges are linked to winter rains followed by increasing temperature and rains in the spring, and to late summer monsoonal rains. Thus, adults of different species are active (for their very short adult lives of 1-2 days) in spring, summer, or both (Waring and Price 1989). The different species in this group are sympatric over a broad area and widely distributed across the Mojave, Sonoran, and Chihuahuan deserts of North America, and up to 10 species having been collected from a single creosote bush (Waring and Price 1989).

In this study, we investigated the phylogenetic relationships of the '*Asphondylia auripila* group' (Gagne and Waring 1990) of cecidomyiid flies, in order to evaluate hypotheses regarding the role of host-plant use in their diversification. First, we used DNA sequence data from one mitochondrial and three nuclear genes to address the hypothesis that the *auripila* group has evolved wholly or in part via *in situ* radiation on *Larrea tridentata.* Second, we analyzed the potential roles of ecology (gall position) and phenology (adult emergence time) in the diversification of this group. Thus, if new species arise in association with changes in gall position, then we expect sister species to exhibit contrasting gall positions. By contrast, if new species arise through phenological separation, then sympatric sister taxa are predicted to be temporally isolated. Alternatively, if neither temporal isolation nor tissue shifts are observed, then new species are more likely to have arisen through divergence resulting from geographic isolation. Lastly, we employed independent contrast analysis to test whether evolutionary shifts in gall position (the host plant part that is galled) are associated with increased rates of change in two ecologically important traits, ovipositor length and wing length.

4.3 Methods

4.3.1 Collection Sites and Methods

We collected *Asphondylia* species associated with *Larrea tridentata* (creosote bush) from sites across southern California, Nevada, Arizona, New Mexico and Texas between March and September 2001-2005. We also collected six Asphondylia species associated with the sympatric host plants *A. atriplicis*, *A. caudicis*, and *A. neomexicana* from saltbush (*Atriplex* spp.), *A. bigeloviabrassicoides* from rabbitbrush (*Chrysothamnus* spp.), *A. websteri* from alfalfa (*Medicago* spp.) and *A.* spp. from snake weed (*Gutterieza* spp.) as putative outgroups. Outgroups were chosen based on previous taxonomic work which identified the saltbush inhabiting *Asphondylia* species as a potential sister group complex to those found on creosote bush based upon shared morphological character states between these two groups (Gagne and Waring 1990). One additional outgroup *A. conglomerata* from a species of saltbush (*Atriplex hamalis*) was obtained from Genbank.

Field-collected galls were transported to the laboratory in an ice-filled cooler where they were kept at room temperature until adults emerged. Following emergence, adults were preserved whole in 20 % dimethyl sulphoxide in a saturated solution of NaCl. Voucher specimens were deposited at the Smithsonian Institution National Museum of Natural History in Washington, DC.

4.3.2 Collection of DNA data

Genomic DNA was isolated using standard phenol chloroform methods (Hillis, Moritz et al. 1996) from single adult midges of either sex. DNA was extracted from as

many individuals for each species as possible (Table 8). We used polymerase chain reactions (PCR) to amplify three nuclear and one mitochondrial gene. A 452 base pair fragment of cytochrome oxidase I (COI) was amplified using primers C1-J-1718 and C1-N-2191 (Simon et al. 1994). A 419 base pair fragment of the internal transcribed spacer region 2 (ITS-2) of nuclear ribosomal DNA (Harris and Crandall 2000) was amplified using primers 5.8sFC and 28s BLD (Simon et al. 1994). A 574 base pair fragment of the Wingless gene (Wg) was amplified using primers 5'wg1 and 3'wg2 (Ober 2002). A 568 base pair fragment of the elongation factor 1 alpha (EF-1∝) gene was amplified using primers EF1aF (AAAATGCCATGGTTCAAAGG) and EF1aR (CGAAATTTGACCTGGATGGT) developed based on an EF-1∝ sequence from *Mayetiola destructor* obtained from Genbank (accession number AF085227). Resulting PCR products were purified using Shrimp Alkaline Phosphatase (SAP) and Exonuclease (EXO), and purified PCR products were used in sequencing reactions with an ABI Prism™ Dye Terminator Cycle sequencing kit.

4.3.3 Phylogenetic Analyses

Sequences were aligned using Clustal (Thompson et al. 1994) and adjusted by eye using Se-Al (Rambaut 1996). Protein coding genes were also checked to ensure that they coded and for stop codons in Se-Al (Rambaut 1996). The best fitting model of sequence evolution was determined for each gene using ModelTest (Posada and Crandall 1998). We also employed MrModeltest 2.2 (Nylander 2004) to identify best models of sequence evolution for each partition for use in Bayesian phylogeny estimation. We first used maximum likelihood (ML) and maximum parsimony (MP) analyses to infer phylogenies for *Asphondylia* species for each gene separately. We employed the heuristic (ML) and

branch and bound (MP) searching features of PAUP 4.0b10 (Swofford 2002). Maximum likelihood trees were also reconstructed using Mr Bayes 3.1.2 (Ronquist and Huelsenbeck 2003). To assess support for recovered nodes, we employed bootstrap replicates (500 for ML, 1000 for MP). We employed the Incongruence Length Test (ILD Test) as implemented in PAUP* (TBR, 1000 replicates) (Huelsenbeck et al. 1996; Swofford et al. 2002), to help evaluate the congruence of the trees inferred from the four different genes. To analyze the combined data, we employed a four partition analysis applying the best fit model of sequence evolution for each partition using Mr.Bayes 3.1.2 (Ronquist and Huelsenbeck 2003).

Evaluation of the monophyly of *Asphondylia* taxa found on *L. tridentata* is complicated by the large number of ingroup taxa (14) relative to putative outgroup taxa (7) in our data set, and size of the genus as a whole (67 Nearctic species, 247 world wide). We used several lines of evidence to test the hypothesis of monophyly. First, we considered MP, ML bootstrap values, and Bayesian posterior probability values from the combined tree, for the nodes that corresponded to monophyly of the *Larrea* taxa (Hillis and Bull 1993). Second, we used Shimodeira-Hasegawa tests and Templeton tests, as implemented in PAUP* (Swofford 2002), to compare the best trees with constraint trees that forced the invasion of the ingroup by one or more outgroup taxa. For example, the best tree was compared to the best constraint tree that did not contain the grouping ((ingroup1, ingroup2, ingroup3, ingroup4)) because one or more outgroup species had invaded the combined ingroup.

4.3.4 Comparative Analyses

We predicted that changes in gall position should be associated with accelerated change in an ecomorphological trait (ovipositor length) related to gall induction, but not in change in wing length, a trait closely indicative of body size (Sokoloff 1966, Norry 1996). To best infer changes in gall position, we utilized Bayesian methods to reconstruct ancestral states for the categorical four-state character gall position (leaf, stem, flower, bud) for each node, using Bayes MultiState (Pagel et al. 2004). This program uses a MCMC approach to sampling phylogenies, and for investigating the parameters of trait evolution, and it calculates a fifth state for the probability that the node does not exist. To calculate the strength of evidence for a shift in gall position at each node we first calculated the probability of no shift across an internode by summing the product of the probability of each state in each node (e.g. p(leaf) node A * p(leaf) node B + p(flower)node A * p(flower) node B) + …, where A and B are the ends of an internode). One minus this probability is a continuous measure of the probability of change for each node, that accounts for phylogenetic uncertainty. To quantify the evolution of our ecomorphological trait (ovipositor length) we optimized this trait, and wing length (a measure of body size), on the combined data Bayesian consensus tree (data from Gagnè and Waring 1990) using McPeek's (1995) contrast method. We then used McPeek's (1995) independent contrast test to determine whether higher rates of change in ovipositor length and wing length occurred along branches associated with ecological shifts (changes in gall position) relative to branches lacking ecological shifts. We tested this hypothesis by regressing a measure of the probability of change at each node with independent contrast values. For this analysis, we used the 'speciational' model

of character evolution, because we assumed that changes in ovipositor morphology take place in association with speciation events rather than continuously over time.

4.4 Results

4.4.1 Dataset

The complete dataset of cytochrome oxidase subunit I (COI), internal transcribed spacer region 2 (ITS-2), wingless (Wg), and elongation factor 1 alpha (EF-1∝) nucleotide sequences for 21 *Asphondylia* species, consisted of 2013 positions (452 COI, 574 Wg, 419 ITS-2, 568 EF-1∝, Table 8). Of 2013 sites, 243 were parsimony informative (118 COI, 47 Wg, 30 ITS-2, 88 EF-1∝). Interspecific pairwise differences within the ingroup ranged from 0.2% - 15.0% for COI, 0.4 - 5.5% for Wg, 0.0% - 2.5% for ITS-2, 0.5%-7.8% EF-1∝. Differences between the ingroup and outgroups were 9.3% COI, 10.5% Wg, 4.8% ITS-2 and 6.6% EF-1∝. Incongruence Length Difference (ILD) tests showed that all-possible combinations of the different gene regions were compatible (p=0.51).

4.4.2 Phylogenies

Figure 8 shows the maximum likelihood trees for each of the four gene regions. For each of the four, MP and ML and Bayesian analyses yielded trees of very similar topology. The grouping of the ingroup taxa into five main clades relative to the outgroup taxa was consistent across all genes except EF-1∝ in which one clade is moved to the base of the tree with the outgroup taxa (Figure 8). The topologies of the best trees for COI and Wg exhibited only minor differences. ITS-2 differed in the placement of one leaf galling taxon (*A. villosa*) and in the placement of *A. florea* and *A. rosetta* at the base

of the tree. EF1-1∝ differed in the invasion of the ingroup by the putative outgroup taxon *A. atriplicis* and in the placement of the stem galling clade (*A. auripila, A. foliosa,* and *A. resinosa*) at the base of the tree with the outgroup taxon. MP, ML and Bayesian analyses of the combined dataset yielded similar topologies (Figure 9).

4.4.3 Evolution of Host Plant Use

All *Asphondylia* species that induce galls on *Larrea tridentata* formed a monophyletic group for both the combined dataset and three of the four datasets separately. Support for the node indicating monophyly of this entire group varied among genes, being strongest in ITS-2, the most highly-conserved gene (100, 100, 100; MP bootstrap, ML bootstrap, and Bayesian a posteriori probabilities respectively), moderate in Wg (61, 73, 95), weakest in COI (-, 65, 99), non-existent in EF-1∝, and intermediate but weak in the combined analyses (60, 56, 62). Monophyly of the *Asphondylia* species on *L. tridentata* was strongly statistically supported for the ITS2 data under MP using Templeton's test (difference in length = 15, $P < 0.001$) and under ML using Shimodaira-Hasegawa's test (difference in –ln L = 48.67, $P < 0.001$). The ML and MP scores for best trees were better than negative constraint trees, but not significantly so as judged by Templeton and Shimodaira-Hasegawa tests for the rest of the datasets and for the combined dataset.

Within the *A. auripila* group, five clades consistently formed strongly supported groups as judged by ML and MP bootstrap support, Bayesian posterior probabilities, Shimodaira-Hasegawa's test, and Templeton's tests. These five clades consisted of three pairs of leaf-galling sister taxa, the clade containing three species which form galls on different plant parts (*Asphondylia rosetta*, *A. florea*, and *A. apicata)*,

and a fifth clade containing four species, three of which form galls on the same plant part but display widely divergent emergence timing (Table 9). These results demonstrate that although support for monophyly of the entire *A. auripila* group of gall midges on *L. tridentata* is not definitive, there is strong evidence for within-host plant speciation within particular clades.

4.4.4 Host-Plant Colonization Sequence

Bayesian ancestral state reconstruction of colonization of different host plant parts yielded several notable inferences (Figure 10). Leaf galling has apparently evolved twice from stem galling ancestors (*A. digitata A. barbata, A. villosa, A. silicula,* and *A. fabalis)* and flower (*A. florea)* and bud galling (*A. apicata*) each evolved once but the order under which these transitions occurred is not clear (Figure 10). In the well-supported clades within this radiation on a single host plant, speciation has apparently occurred in association with shifts to new plant parts three times, and in association with retention of the same host plant part five times.

4.4.5 Ecological Adaptation to Specific Plant Parts

4.4.5.1 Evolution of phenology

Most *Asphondylia* species (10 of 14 sampled) found on *Larrea tridentata* are bivoltine, with adults found in both spring and summer. The remaining four species are univoltine, being found as adults in only the spring, winter, or summer as follows: March-May (*A. foliosa*), August-September (*A. rosetta*, *A. auripila*), and December-February (*A. resinosa*) (Figure 10). If new species arise through phenological separation, then sympatric sister taxa are expected to be temporally isolated. Among well supported

clades, sister-taxa comparisons for phenology (Figure 10) show two main patterns: (1) bivoltine sister taxa emerge at the same times (*A. barbata* and *A. villosa; A. silicula* and *A. fabalis*; *A. clavata* and *A. pilosa*; *A. apicata* and *A. florea*) and (2) three of the four univoltine taxa that are phenologically isolated are members of the same clade (*A. resinosa*, *A. auripila*, and *A. foliosa*) and within this clade there is a reversal to bi-voltinism (*A. digitata*).

4.4.5.2 *Evolution of ecomorphology*

Our Bayesian extension of McPeek's (1995) contrast analysis indicate that ovipositor length underwent higher rates of change along branches where shifts to new plant parts were inferred to be more probable than where shifts were not inferred (R^2=0.26, F= 5.054, DF=11, $p < 0.05$). By contrast, there was no difference in rates of change in wing length in relation to shifts in plant part vs. retention of the same plant part (R^2=0.13, F= 2.802, DF=11, p =0.13).

4.5 Discussion

Four of five phylogenies (each gene and combined), and SH and Templeton tests for ITS2, support the hypothesis derived from morphological data (Gagné and Waring 1989) that the *A. auripila* group has radiated *in situ* on *Larrea tridentata*. The tree from the EF-1∝ data did not support the hypothesis of monophyly for the *A. auripila* group as a whole. However, SH and Templeton's tests show this tree is not significantly better than a tree constraining the ingroup (*Asphondylia auripila* group) to be monophyletic. The contrasting results from different genes, and the non-significant SH and Templetons

tests, suggest that based on the currently available evidence, support for monophyly of the *A. auripila* group as a whole remains equivocal.

Despite this uncertainty regarding monophyly of the *A. auripila* group as a whole, two lines of evidence strongly support the monophyly of multiple clades within this group. First, Bayesian posterior probabilities, maximum likelihood, and maximum parsimony bootstap values indicate strong support for five clades, and some of the sister-species in these clades are very closely related (e.g. *A. villosa* and *A. barbata* differ by only 1.3 percent at COI). Second, Shimodaira-Hasegawa and Templetons tests significantly support the hypotheses of monophyly of these clades (Figure 10). Thus, even if the entire *A. auripila* group is not monophyletic, it comprises multiple lineages that show strong evidence for monophyly, which indicates that this group is characterized by a notable degree of within host-plant speciation. Hypotheses regarding monophyly of this clade, and the lineages within it, could be tested further via sequencing of additional *Asphondylia* species from the North American deserts.

4.5.1 Potential Mechanisms of Within Host-plant Speciation

Shifts to a new host plant are usually accompanied by adaptations to markedly different plant characteristics, such as plant morphology, chemistry and phenology (Jaenike 1989; Jaenike 1990; Becerra and Venable 1999; Cook, Rokas et al. 2002). By contrast, shifts within a host plant may not require such substantial evolutionary change. Other barriers, such as high rates of gene flow, likely inhibit speciation via ecological shifts within a host plant (Ferdy et al. 2002). In *Asphondylia* midges, there are several possible geographic modes and mechanisms of speciation within a single host plant, each

of which could result in the partitioning of the plant into a number of finely-divided niches.

4.5.1.1 Divergence under sympatry

Changes in diapause timing could result in sympatric populations shifting in time to exploit the same or a new part of a host plant at a different point in time, effectively generating reproductive isolation. Thus, three species of stem galling *Asphondylia* midges on *Larrea tridentata* (*A. auripila, A. foliosa, A. resinosa*) in a well-supported clade are phenologically separated from one another (Figure 10). The emergence timing of these species corresponds to the timing of plant growth associated with rains in winter (*A. resinosa*), spring (*A. foliosa*) and summer (*A. auripila*). The emergence timing of other members of the *Asphondylia auripila* group show no seasonal isolation between sister taxa, although they may be phenologically isolated on a finer scale (within a season), given the short life spans and weak flight abilities of adult flies (Jones et al. 1983; Gagné 1989). This hypothesis could be tested by monitoring the emergence timing of bivoltine sister taxa such as *A. barbata* and *A. villosa*.

Phenology has been shown to be important in mediating reductions in gene flow leading to speciation or host race formation in many other insect taxa, including *Rhagoletis* flies (Feder and Filchak 1999), *Eurosta* flies (Craig, Itami et al. 1993), *Enchenopa* treehoppers (Wood, Olmstead et al. 1990), *Magicicada* cicadas (Wood, Olmstead et al. 1990; Cooley, Simon et al. 2003), and *Blepharoneura* flies (Condon and Steck 1997). These parallel patterns suggest that temporal isolation may be an important process favoring speciation in phytophagous insects.

Phenological divergence may be facilitated by shifts to competition-free space, in that the insects that have shifted to a new plant part are expected to be released from the strong competition that typifies many gall-inducing species (Denno 1995, Inbar 1998, Craig et al. 2000). The prolonged diapause of the gall midge *Dasineura rachiphaga* is thought to be a mechanism that evolved in the context of selection for reduced intraspecific competition for limiting oviposition sites (Prévost 1990). Similarly, Cook et al. (2002) showed that speciation of *Andricus* gall wasps is more commonly associated with shifts to a novel part of the same host plant than with shifts between different host plant species, and they suggested that intraspecific competition for oviposition sites has facilitated within-host divergence. In *Chiastocheta* flies inhabiting *Trollius* species, Després et al. (2002) demonstrated that diversification has involved both shifting hosts and radiation within a host, and the within-host diversification may be a result of competition for oviposition or feedings sites favoring temporal shifts in oviposition timing and shifts to different larval food resources (Ferdy, Després et al. 2002).

The proximate mechanism of sympatric shifts in host-plant parts may involve a combination of mistakes in oviposition site, and variation in the developmental schedules of different plant parts. Insects sometimes lay eggs on unfamiliar host plants or host plant parts; such ovipositional mistakes have been documented for Lepidoptera (Feldman and Haber 1998), Coleoptera (Fox et al. in press), and Diptera (Gratton and Welter 1998), including many Cecidomyiidae (Larsson and Strong 1992, Larsson and Ekbom 1995). When a female oviposits on a plant tissue type other than her natal type (i.e. flower instead of leaf), the eggs in the new tissue type may break diapause later or earlier as a result of differences in the developmental schedule of the different plant tissue types

(Linkosalo 2000, Mahoro 2002), and this may translate to the temporal isolation of adults. This hypothesis could be tested with the *Asphondylia* midges on *Larrea tridentata* by enforcing oviposition on non-natal host-plant parts (i.e. leaf – stem) and recording changes in emergence timing.

4.5.1.2 Divergence under allopatry

Colonization of a new plant part could also occur in an allopatric population, resulting in a single species inducing galls on multiple parts of a single host plant. The ability to gall the original part of the host plant may, in theory, be subsequently lost, or the colonizing species may go locally extinct, and differentiation could then occur due to drift and selection in allopatry. Upon recontact, we are left with two sympatric species utilizing different niches on the same host plant. Speciation on the same plant part could also result from allopatric isolation. In this scenario reproductive isolation and ecological divergence develop as a product of isolation through both selection resulting from different ecological conditions (climate, plant genotype, parasitoids, and composition of the galling community) and differentiation due to genetic drift. Upon recontact we have two ecologically diverged species (e.g. phenologically isolated) on the same plant part. In a third scenario, reproductive isolation could develop in allopatry purely due to genetic drift, and ecological divergence of the resulting species occurs as a result of subsequent interspecific competition. The host plant of the *Asphondylia auripila* group is the dominant shrub throughout an immense area, the southwestern deserts of North America (Hunter et al. 2001). *Larrea tridentata* was isolated in refugia during the major North American glaciations (Hunter, Betancourt et al. 2001), and speciation may have occurred in this manner in refugia during glacial periods. However, by this hypothesis it is not

clear why the ability to gall the original plant part would be lost, or why such progenitor populations would go extinct; moreover most of the radiation on *L. tridentata* appears to be considerably older than the glaciation cycles starting in the Pleistocene. This allopatry hypothesis could be addressed further through comparative phylogeographic analyses of sister-taxa inducing galls on different plant parts.

4.5.2 Ecological Adaptation to Specific Plant Parts

Adaptive changes in insect morphological characters following host shifts have been documented only rarely, despite the central importance of morphological adaptations in insect diversification (Moran 1986; Carroll et al. 1997; Groman and Pellmyr 2000). In this study, we have documented adaptive changes in an ecologically-important morphological character, ovipositor length, within the context of radiation on a single host-plant species. Our independent contrast analyses, which account for both uncertainty in the phylogeny and uncertainty in the reconstructions of ancestral galling position states, demonstrate that *Asphondylia* species inhabiting *Larrea tridentata* show substantially larger changes in ovipositor length following ecological shifts (shifts to new parts of a host plant) relative to the amount of change when no ecological shift has taken place. By contrast, wing length, a trait not predicted to be adaptive in the context of exploitation of different plant parts, shows no significant relation with ecological shifts. The finding that ovipositor length changed more than wing length in response to ecological shifts suggests that selection for host-plant part associated morphological differences is driving changes in *Asphondylia* ovipositor lengths.

The morphological basis of adaptation to different host plant parts in these species is simple: *Asphondylia* species inhabiting different parts of *Larrea tridentata* deposit

their eggs into strikingly different tissue types (stems, leaves, buds, and flowers) that differ markedly in hardness, thickness, and depth to plant vasculature. Thus, the shorter ovipositor of leaf galling species may facilitate the placement of eggs in thinner softer leaf tissue, whereas longer ovipositors of stem, bud, and flower galling species allow egg placement deeper into host plant tissues. These findings suggest that strong divergent selection on ovipositor length accompanies evolutionary shifts in host-plant part, which would be expected to drive post-zygotic isolation; this hypothesis could be tested further via measuring oviposition depths in different plant tissues, and through experimental manipulation of oviposition sites.

4.5.3 Conclusions

Our study provides strong evidence that some clades of *Asphondylia* gall midges have radiated *in situ* on their host plant *Larrea tridentata*. This diversification was apparently driven by the ability of these insects to partition the plant ecologically, via two mechanisms that facilitate the evolution of reproductive isolation: shifts to new plant parts and changes in phenology. Evidence from other host-specific phytophagous insects that can utilize different parts of the same plant species (e.g., Condon and Steck 1997; Cook, Rokas et al. 2002; Després, Pettex et al. 2002), and from host-specific parasites (e.g., Simková et al. 2004), suggests that within-host ecological divergence may be a common mechanism of speciation that promotes the extraordinarily high species diversity found in many groups of parasites and plant-feeding insects.

4.6 Acknowledgments

We are especially grateful to A. Mooers, N. Moran, C.E. Parent, P. Nosil and 4 anonymous reviewers for advice and comments that greatly improved this manuscript, and to the SFU FAB* lab for discussion. We are also very grateful to N. Moran for laboratory space and support in Arizona. H. Kucera and F. Joy assisted with field data collection. We thank R. Gagné and the Smithsonian Institution National Museum of Natural History for providing samples of museum preserved cecidomyiid species. This research was funded by the Natural Sciences and Engineering Research Council of Canada and by the Society for Systematic Biology.

4.7 Literature Cited

Abrahamson, W. G., M. D. Eubanks, C. P. Blair, and A. V. Whipple. 2001. Gall flies, inquilines, and goldenrods: A model for host-race formation and sympatric speciation. Am. Zool. 41:928-938.

Becerra, J. X., and D. L. Venable. 1999. Macroevolution of insect-plant associations: The relevance of host biogeography to host affiliation. Proc. Natl. Acad. Sci. U.S.A. 96:12626-12631.

Berlocher, S. H. 2000. Radiation and divergence in the *Rhagoletis pomonella* species group: Inferences from allozymes. Evolution 54:543-557.

Caillaud, M. C., and S. Via. 2000. Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. Am. Nat. 156:606-621.

Carroll, S. P., H. Dingle, and S. P. Klassen. 1997. Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. Evolution 51:1182-1188

Condon, M. A., and G. J. Steck. 1997. Evolution of host use in fruit flies of the genus *Blepharoneura* (Diptera: Tephritidae): cryptic species on sexually dimorphic host plants. Biol. J. Linn. Soc. 60:443-466.

Cook, R. J., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gallwasps. Evolution 56:1821-1830.

Cooley, J. R., C. Simon, and D. C. Marshall. 2003. Temporal separation and speciation in periodical cicadas. Bioscience 53:151-157.

Craig, T. P., J. K. Itami, W. G. Abrahamson, and J. D. Horner. 1993. Behavioral evidence for host-race formation in *Eurosta solidaginis*. Evolution 47:1696-1710.

Craig, T. P., J. K. Itami, C. Shantz, W. G. Abrahamson, J. D. Horner, and J. V. Craig. 2000. The influence of host plant variation and intraspecific competition on oviposition preference and offspring performance in the host races of *Eurosta solidaginis*. Ecol. Entomol. 25:7-18.

Craig, T. P., J. D. Horner, and J. K. Itami. 2001. Genetics, experience, and host-plant preference in *Eurosta solidaginis*: Implications for host shifts and speciation. Evolution 55:773-782.

Crespi, B. J., D. A. Carmean, and T. W. Chapman. 1997. Ecology and evolution of galling thrips and their allies. Annu. Rev. Entomol. 42:51-71.

Crespi, B., and M. Worobey. 1998. Comparative analysis of gall morphology in Australian gall thrips: The evolution of extended phenotypes. Evolution 52:1686-1696.

Denno, R. F., M. S. McClure, and J. R. Ott. 1995. Interspecific interactions in phytophagous insects - competition reexamined and resurrected. Annu. Rev. Entomol. 40:297-331.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. 2002. Speciation in the globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258-68.

Dorchin, N., A. Freidberg, and O. Mokady. 2004. Phylogeny of the *Baldratiina* (Diptera : Cecidomyiidae) inferred from morphological, ecological and molecular data sources, and evolutionary patterns in plant-galler relationships. Mol. Phylogenet. and Evol. 30:503-515

Drès, M., and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. Lond. Ser. B 357:471-492.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and plants: a study in coevolution. Evolution 18:586-608.

Emelianov, I., M. Drès, W. Baltensweiler, and J. Mallet. 2001. Host-induced assortative mating in host races of the larch budmoth. Evolution 55: 2002-10.

Farrell, B. D., and C. Mitter. 1994. Adaptive radiation in insects and plants - time and opportunity. Am. Zool. 34:57-69.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja. 2003. Allopatric genetic origins

for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. U.S.A. 100:10314-10319.

Feder, J. L., and K. E. Filchak. 1999. It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. Entomol. Exp. Appl. 91:211-225.

Feder, J. L., S. B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. Proc. Natl. Acad. Sci. U.S.A 91:7990-4.

Feldman, T. S., and W. A. Haber. 1998. Oviposition behavior, host plant use, and diet breadth of *Anthanassa* butterflies (Lepidoptera : Nymphalidae) using plants in the Acanthaceae in a Costa Rican community. Fla. Entomol. 81:396-406.

Ferdy, J. B., L. Després, and B. Godelle. 2002. Evolution of mutualism between globeflowers and their pollinating flies. J. Theor. Biol. 217:219-234.

Fox, C.W., R.C. Stillwell, A.R. Amarillo-s, M.E. Czesak, and F.J. Messina. Genetic architecture of population differences in oviposition behaviour of the seed beetle *Callosobruchus maculatus*. J. Evol. Biol. *In Press.*

Funk, D. J., K. E. Filchak, and J. L. Feder. 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116:251-267.

Futuyma, D. J., M. C. Keese, and D. J. Funk. 1995. Genetic constraints on macroevolution - the evolution of host affiliation in the leaf beetle genus *Ophraella*. Evolution 49:797-809.

Gagné, R. J. 1989. The Plant-feeding Gall Midges of North America. Cornell University Press, Ithaca.

Gagné, R. J., and G. L. Waring. 1990. The *Asphondylia* (Cecidomyiidae: Diptera) of creosote bush (*Larrea tridentata*) in North America. Proc. Entomol. Soc. Wash. 92:649-671.

Gagné, R. J., and P.E. Boldt. 1991. *Asphondylia* (Diptera: Cecidomyiidae) does not reflect the disjunct distribution of *Larrea* (Zygophyllaceae). Proc. Entomol. Soc. Wash. 93:509-510.

Gratton, C., and S. C. Welter. 1998. Oviposition preference and larval performance of *Liriomyza helianthi* (Diptera : Agromyzidae) on normal and novel host plants. Environ. Entomol. 27:926-935.

Groman, J. D., and O. Pellmyr. 2000. Rapid evolution and specialization following host colonization in a yucca moth. J. Evol. Biol. 13:223-236.

Harris, D. J., and K. A. Crandall. 2000. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda: Cambaridae): Implications for phylogenetic and microsatellite studies. Mol. Biol. Evol. 17:284-291.

Hawkins, B. A., R. D. Goeden, and R. J. Gagné. 1986. Ecology and taxonomy of the *Asphondylia* spp. (Diptera: Cecidomyiidae) forming galls on *Atriplex* spp. (Chenopodiaceae) in Southern California. Entomography 4:55-107.

Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42:182-192.

Hillis, D. M., C. Moritz, and B. K. Mable. 1996. Molecular Systematics (2nd Edition). Sinauer and Associates.

Huelsenbeck, J. P., and J. J. Bull. 1996. A likelihood ratio test to detect conflicting phylogenetic signal. Syst. Biol. 45:92-98.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-5.

Hunter, K. L., J. L. Betancourt, B. R. Riddle, T. R. Van Devender, K. L. Cole, and W. G. Spaulding. 2001. Ploidy race distributions since the last glacial maximum in the North American desert shrub, *Larrea tridentata*. Glob. Ecol. and Biogeogr. 10:521-533.

Inbar, M., M. Wink, and D. Wool. The evolution of host plant manipulation by insects: molecular and ecological evidence from gall-forming aphids on *Pistacia*. Mol. Phylogenet. Evol. *In press*.

Jaenike, J. 1989. Genetics of butterfly hostplant associations. Trends Ecol. Evol. 4:34-34.

Jaenike, J. 1990. Host specialization in phytophagous insects. Annu. Rev. Ecol. Syst. 21:243-273.

Jermy, T. 1984. Evolution of insect/host plant relationships. Am. Nat. 124:609-630.

Jones, R. G., R. J. Gagné, and W. F. Barr. 1983. Biology and taxonomy of the *Rhopalomyia* gall midges (Diptera: Cecidomyiidae) of *Artemesia tridentata* Nutall (Compositae) in Idaho. Contrib. Am. Entomol. Instit. 21:1-76.

Larsson, S., and D. R. Strong. 1992. Oviposition choice and larval survival of *Dasineura marginemtorquens* (Diptera, Cecidomyiidae) on resistant and susceptible *Salix viminalis*. Ecol. Entomol. 17:227-232.

Larsson, S., and B. Ekbom. 1995. Oviposition mistakes in herbivorous insects - confusion or a step towards a new host? Oikos 72:155-160.

Linkosalo, T. 2000. Mutual regularity of spring phenology of some boreal tree species: predicting with other species and phenological models. Can. J. For. Res. 30:667-673.

Mahoro, S. 2002. Individual flowering schedule, fruit set, and flower and seed predation in *Vaccinium hirtum* Thunb. (Ericaceae). Can. J. Bot. 80:82-92.

Mardulyn, P., M. C. Milinkovitch, and J. M. Pasteels. 1997. Phylogenetic analyses of DNA and allozyme data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. Syst. Biol. 46:722-747.

McPeek, M. A. 1995. Testing hypotheses about evolutionary change on single branches of a phylogeny using evolutionary contrasts. Am. Nat. 145:686-703.

Moran, N. A. 1986. Morphological adaptation to host plants in *Uroleucon* (Homoptera, Aphididae). Evolution 40:1044-1050.

Norry, F. M., and J. C. Vilardi. 1996. Size-related sexual selection and yeast diet in *Drosophila buzzatii* (Diptera: Drosophilidae). J. Insect Behav. 9:329-338.

Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nyman, T., A. Widmer, and H. Roininen. 2000. Evolution of gall morphology and host-plant relationships in willow-feeding sawflies (Hymenoptera : Tenthredinidae). Evolution 54:526-533.

Ober, K. 2003. Phylogenetic relationships of the carabid subfamily Harpalinae (Coleoptera) based on molecular sequence data. Mol. Phylogenet. Evol. 26:334-336.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral states on phylogenies. Systematic Biology 53(5):673-684.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-8.

Prevost, Y. H. 1990. Spruce cone axis midge, *Dasineura rachiphaga* Tripp (Diptera, Cecidomyiidae), in cones of black spruce, *Picea mariana* (Mill) Bsp. Can. Entomol. 122:441-447.

Price, P.W., G.W. Fernandes, A.C.F. Lara, J. Brawn, H. Barrios, M.G. Wright, S.P. Ribeiro, and N. Rothcliff. 1998. Global patterns in local number of insect galling species. Journal of Biogeography 25: 581-591.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk/.

Ronquist, F., and J. Liljeblad. 2001. Evolution of the gall wasp-host plant association. Evolution 55:2503-22.

Simkova, A., S. Morand, E. Jobet, M. Gelnar, and O. Verneau. 2004. Molecular phylogeny of congeneric monogenean parasites (Dactylogyrus): a case of intrahost speciation. Evolution 58:1001-1018.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook 1994, Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annu. Rev. Entomol. 87: 651-701.

Sokoloff, A. 1966. Morphological variation in natural and experimental populations of *Drosophila pseudoobscura* and D. persimilis. Evolution. 20:49–71.

Stone, G. N., and J. M. Cook. 1998. The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. Proc. Roy. Soc. London B 265:979-88.

Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J. N. 1994. The Coevolutionary Process. University of Chicago Press.

Thompson, J. D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22: 4673-4680.

Waring, G. L., and P. W. Price. 1989. Parasitoid pressure and the radiation of a gallforming group (Cecidomyiidae: *Asphondylia* spp.) on creosote bush (*Larrea tridentata*). Oecologia 79:293-299.

Wood, T. K., K. L. Olmstead, and S. I. Guttman. 1990. Insect phenology mediated by host-plant water relations. Evolution 44:629-636.

Table 8 Number of sequences obtained per species per gene (see also Table 9).

Species	Host Plant	COI	ITS-2	Wg	EF-1
A. apicata	*Larrea tridentata*	2	1	0	1
A. rosetta	*Larrea tridentata*	2	2	2	2
A. florea	*Larrea tridentata*	2	2	2	2
A. auripila	*Larrea tridentata*	2	2	2	2
A. foliosa	*Larrea tridentata*	2	2	2	2
A. resinosa	*Larrea tridentata*	2	2	2	2
A. barbata	*Larrea tridentata*	2	2	2	2
A. clavata	*Larrea tridentata*	2	2	2	2
A. fabalis	*Larrea tridentata*	2	1	1	1
A. pilosa	*Larrea tridentata*	2	2	2	2
A. silicula	*Larrea tridentata*	2	1	0	1
A. villosa	*Larrea tridentata*	2	2	2	1
A. digitata	*Larrea tridentata*	1	0	0	0
A. bullata	*Larrea tridentata*	2	0	0	1
A. caudicis	*Atriplex spp.*	1	1	1	0
A. atriplicis	*Atriplex spp.*	1	1	1	1
A. neomexicana	*Atriplex spp.*	1	0	0	0
A. bigeloviabrassicoides	*Chysothamnus spp.*	1	0	1	1
A. spp.	*Gutterizia spp.*	1	0	0	1
A. websteri	*Medicago spp.*	1	0	0	0
A. conglomerata	*Atriplex spp.*	1	0	0	0

Table 9 Summary of support for monophyly of clades within the *A. auripila* group. Support for each of the 5 clades is provided: ML = Maximum Likelihood bootstrap support; MP = Maximum Parsimony bootstrap support; MCMCMC = Bayesian posterior probability; SH = significance for the SH test; Templetons Test = significance level for Templetons test. A * indicates a clade for which there is good support for monophyly; n.s.= not significant.

***A. auripila* Supported Clade**	**MP**	**ML**	**MCMCMC**	**SH Test**	**Templetons Test**
*A. clavata, A. pilosa**	100	100	99	P<0.05	P<0.001
*A. silicula, A. fabalis**	100	100	100	P<0.05	P<0.001
*A. barbata, A. villosa**	100	100	99	P<0.05	n.s.
*A. rosetta, A. florea, A. rosetta**	100	100	96	P<0.05	P<0.05
*A. resinosa, A. auripila, A. foliosa, A. digitata**	88	88	88	P<0.05	P<0.05

Table 10 Collection locations for *Asphondylia* samples used in this study. n/a refers to coordinates not available.

Species	Location	Collection Location	
		Latitude	Longitude
A. apicata	Arizona	32.85419	-112.76898
A. apicata	Arizona	32.85419	-112.76898
A. rosetta	Arizona	33.66552	-114.00259
A. rosetta	Arizona	35.62776	-114.42500
A. florea	Arizona	32.10646	-110.02626
A. florea	Arizona	32.04849	-111.39339
A. auripila	New Mexico	32.22744	-108.95309
A. auripila	Arizona	32.19672	-112.46421
A. foliosa	Arizona	33.43421	-112.58794
A. foliosa	Arizona	32.19672	-112.46421
A. resinosa	Arizona	33.79714	-112.13309
A. resinosa	Arizona	34.05390	-112.14478
A. barbata	Arizona	32.17640	-112.26275
A. barbata	Arizona	34.61367	-111.86295
A. clavata	Arizona	32.04849	-111.39339
A. clavata	Arizona	32.08436	-110.81089
A. fabalis	Arizona	33.40855	-112.39408
A. fabalis	Arizona	33.40855	-112.39408
A. pilosa	Arizona	33.79716	-112.13789
A. pilosa	Arizona	32.46565	-112.87441
A. silicula	Texas	31.06663	-104.21716
A. silicula	Arizona	32.27415	-110.95036
A. villosa	Arizona	31.96300	-110.80246
A. villosa	Arizona	31.96300	-110.80246
A. digitata	Arizona	32.06105	-110.77532
A. bullata	Texas	31.06663	-104.21716
A. bullata	Texas	31.06663	-104.21716
A. caudicis	California	34.92229	-117.27702
A. atriplicis	Arizona	32.75440	-110.64789
A. neomexicana	Arizona	32.75440	-110.64789
A. bigeloviabrassicoides	British Columbia	49.23960	-119.40010
A. spp.	California	32.63629	-116.11862
A. websteri	Arizona	n/a	n/a
A. conglomerata	Israel	n/a	n/a

Figure 8 Maximum likelihood phylogenies for (A) COI, (B) Wingless (C) ITS-2 and (D) EF-1∞. Maximum parsimony, maximum likelihood, and Bayesian support values are shown for each node. Branch lengths are proportional to the inferred number of substitutions per site.

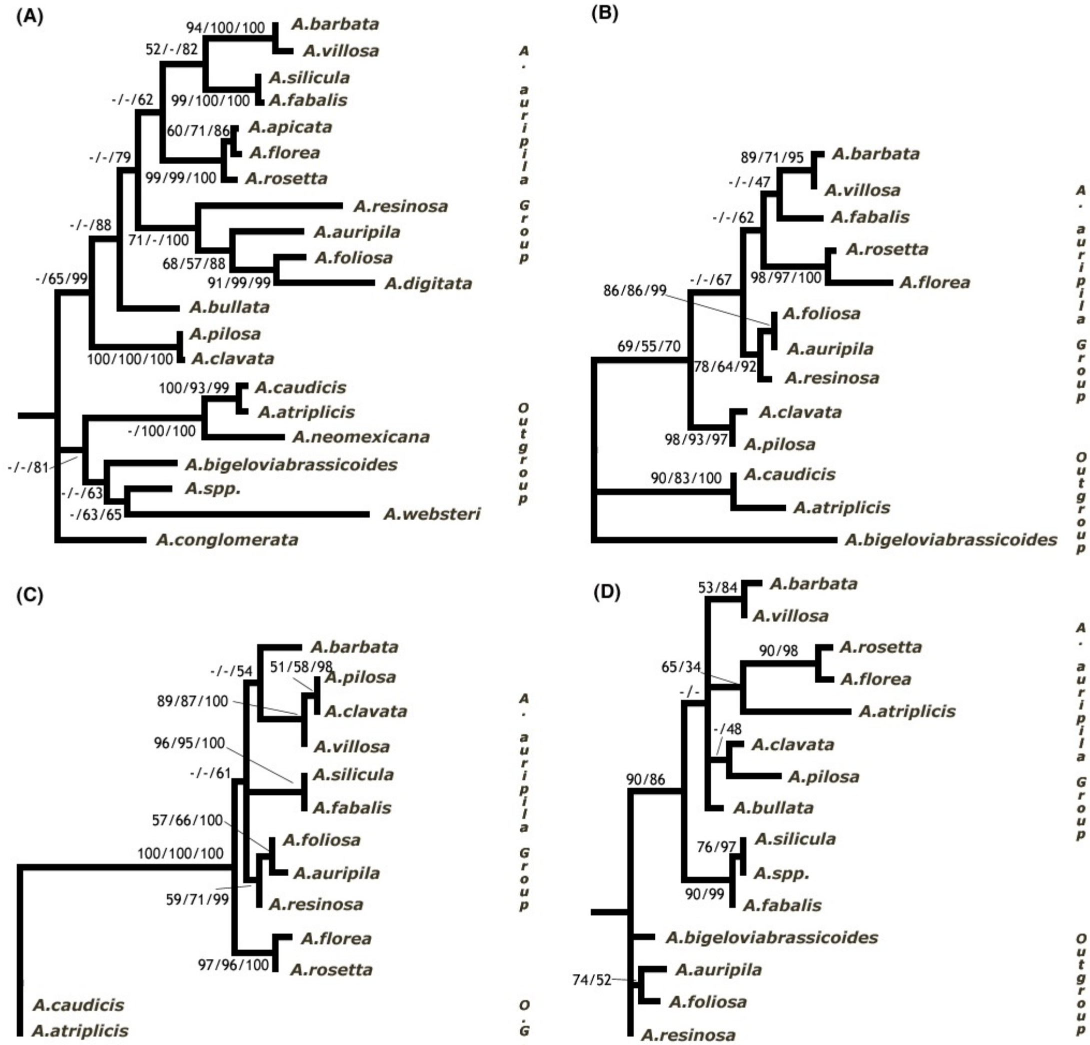


Figure 9 Phylogeny of *Asphondylia auripila* group and outgroups according to a 4-partition Bayesian phylogenetic analysis using a separate substitution model for each gene. Numbers above branches are MP bootstrap, ML bootstrap, and Bayesian posterior probabilities. Host genera are delineated at the tips.

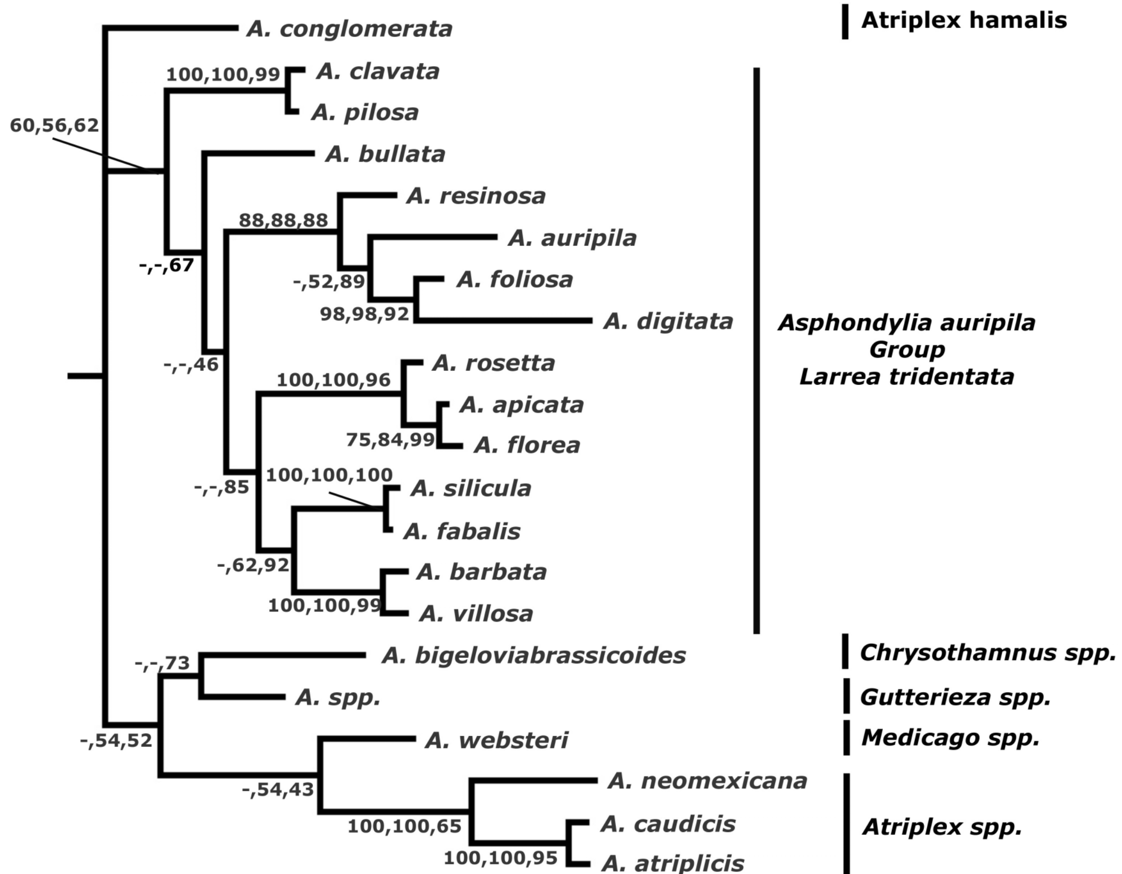


Figure 10 Phylogeny of *Asphondylia auripila group* based on combined dataset with ancestral gall position reconstruction by Bayesian methods. Drawings of galls for each species, gall position, and the phenology of adult emergence is provided for each species.

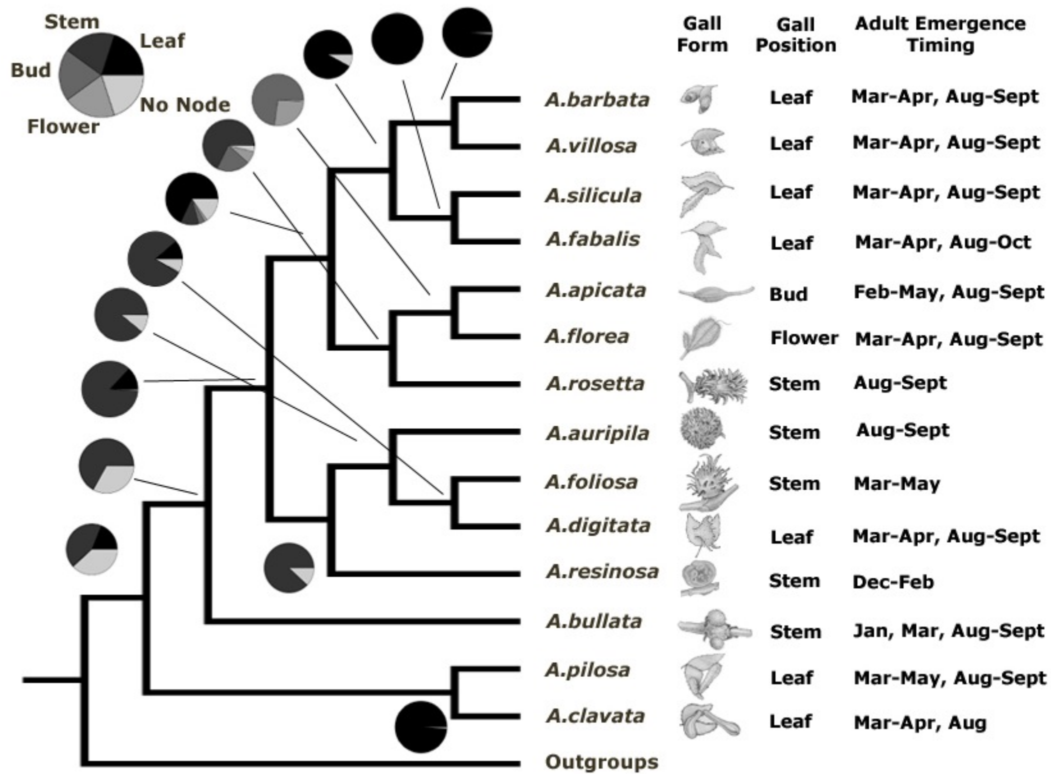
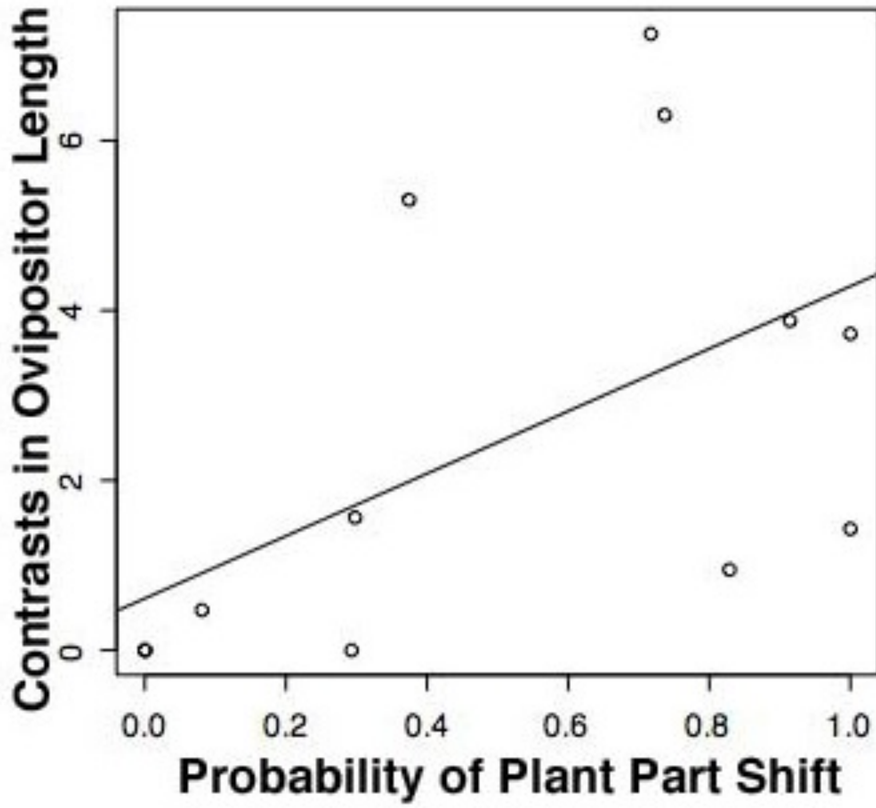


Figure 11 Phylogenetically independent contrast values for *Asphondylia* ovipositor lengths calculated for each node plotted versus an index of the probability of change in host-plant part usage from one node to the next.



Chapter 5: Historical gene flow and population demographic history during adaptive radiation of gall-inducing insects

5.1 Abstract

The colonization of a new niche and persistence in it during adaptive radiation involves complex interactions between natural selection, population size and gene flow. We evaluated the roles of ecology, historical gene flow, and historical population size in the adaptive radiation of a monophyletic group of gall-forming midges with overlapping ranges, all of whom parasitize the same host plant species. Mitochondrial and nuclear DNA markers from two pairs of closely-related species with contrasting forms of ecological divergence were used to identify and map ancestral alleles and to estimate levels of historical gene flow since divergence. Putatively ancestral alleles show substantial geographic overlap for species in each pair, and maximum likelihood reconstructions of ancestral locations show common ancestral locations for members of each pair, consistent with divergence of each pair in geographical proximity. Multilocus coalescent and phylogenetic analyses support a hypothesis of ancient gene flow, and large ancestral population sizes, for the *A. auripila* and *A. foliosa* pair, species that displays divergence in life history timing, but no ancient gene flow, and small ancestral populations, for *A. florea* and *A. rosetta*, the pair of species that displays divergence in plant-part use. Our analyses suggest that local adaptation and population divergence is prevented by even low levels of gene flow in small populations, whereas local adaptation

in populations of relatively larger effective size may be tolerant to low levels of gene flow during population divergence, a pattern consistent with the neutral theory of molecular evolution and theories of ecological speciation.

Keywords: Divergence population genetics, isolation with migration, *Asphondylia*, within-host speciation, plant-part use, phenology.

5.2 Introduction

The ecological contexts in which gene flow, population demographic history, and natural selection promote or retard population divergence during the evolution of reproductive isolation remain a focus of intense interest in evolutionary biology. Speciation in many plant-feeding insect groups involves colonization of a new host-plant species followed by the development of reproductive isolation, in part as a consequence of adaptation to the novel host-plant (Farrell and Mitter 1994). Despite a notable pattern of host-plant shifts coincident with speciation in phytophagous insects, reproductive isolation in an increasing number of plant-feeding insect species has been shown to evolve in the absence of shifts to new species of host plants (Condon et al. 2008, Marstellar et al. 2009, Cook et al. 2002, Sachet et al. 2006, Joy and Crespi 2007). These cases of speciation without host-plant shifts are of particular interest because species divergence may have occurred in varying population demographic contexts with some level of contact and gene flow between diverging populations. High rates of gene flow are expected to hinder local adaptation and divergence, while low rates of gene flow between diverging species are thought to have little effect on local adaptation and divergence (reviewed in Guillaume and Whitlock 2007). Population genetics theory suggests that natural selection operates more efficiently in colonizing populations of large

effective size (Tufto 2001, Lenormand 2002, Rosenblum et al 2007). However, empirical examples inferring the degree of historical gene flow and population demography history during divergence remain scarce (c.f. Coyne and Orr 2004, Knowles and Carstens 2007, Rosenblum et al. 2007, Nadachowska and Babik 2009), and such studies have yet to connect analyses of historical gene flow and demography with ecological contexts of divergence.

Phylogenetic and ecological studies of gall-inducing flies of the *Asphondylia auripila* (Diptera: Cecidomyiidae) group have revealed broad sympatry and niche similarity in a group of species whose members all induce galls on creosote bush (*Larrea tridentata*; (Waring and Price 1989). These flies colonized creosote bush following its arrival in North America from South American progenitors between 8.4 and 4.2 million years ago (Gagne and Boldt 1991, Lia et al. 2001). The species-level phylogeny of this group and related species has revealed a considerable degree of within host-plant radiation, which may have occurred, in part, in non-allopatric settings (Joy and Crespi 2007).

Here, we investigate the conditions of divergence of two pairs of closely-related species of gall-inducing midges (Joy and Crespi 2007). The first pair of species utilizes the same plant part (the stem) but differs in life history timing, with *A. foliosa* exploiting spring plant growth and *A. auripila* utilizing summer growth. By contrast, the second pair of species displays ecological divergence in plant-part use: *A. rosetta* induces galls on the stem and *A. florea* induces galls on the flower, but they overlap broadly in the timing of their emergence from these galls. To the extent that ecological factors drive speciation in this group, we expect pairs of species exhibiting ecological differentiation

(such as *A. rosetta* and *A. florea*) to show stronger, more-rapid divergence than species pairs that differ in life-history timing but are otherwise ecological similar (such as *A. auripila* and *A. foliosa*).

We used a combination of phylogenetic and coalescent based methods to infer the evolutionary history of these species and to evaluate the roles of gene flow and demography in their diversification. First, we used mitochondrial and nuclear sequence data to evaluate the hypothesis that a combined intra-specific and inter-specific genealogy will show that each described species predominantly forms a monophyletic group, indicating that the species based on life history and morphological differences are real units that do not currently interbreed. Second, we evaluated the hypothesis that species within each pair diverged in association with geographical separation. To this end, we constructed haplotype networks to identify relatively-ancient alleles in each species, and we assessed whether these alleles were geographically separated; we also utilized a recently-developed maximum likelihood framework to estimate the ancestral location of each species (Lemmon and Lemmon 2008). If divergence was facilitated by geographic separation, then ancient alleles within each pair are expected to show geographical separation (Hickerson and Cunningham 2005). By contrast, if divergence occurred in the absence of geographical separation, then ancient alleles within each pair are expected to show geographic association. Third, we estimated the degree of gene flow between the two species pairs since divergence, using coalescent simulations (Beerli and Felsenstein 1999, Hey and Neilsen 2004). In these analyses, a complete lack of past migration would support scenarios of divergence without gene flow, while non-zero past migration rates would support a hypothesis of speciation involving episodes of gene flow.

Fourth, we characterized the relative demographic contexts of divergence in each species pair, using mitochondrial and nuclear DNA sequence data in a coalescent framework. Speciation events associated with plant-part shifts (often a result of ovipositional accidents; Price 2005) are expected to involve small founding population sizes. By contrast, the demographic context of lineages which diverge through temporal shifts may be either small (in the case, for example, of mutation in few individuals) or large (in the case of a large population which splits gradually as a result of change in the temporal environment).

5.3 Methods

5.3.1 Collection Sites and Methods

Asphondylia species associated with *Larrea tridentata (*creosote bush*)* were collected between March 2003 and August 2006 from sites across southern California, Nevada, Arizona, New Mexico and Texas. Members of this group differ in several ecologically-important characteristics such as gall morphology, gall position, and ovipositor length (Gagné and Waring 1990). The development and emergence timing of these midges is linked to winter rains followed by increasing temperature, spring rains, and/or late-summer monsoonal rains (Gagné and Waring 1990). The short-lived adults of the different species are active in winter, spring, late summer, or a combination of these periods (Waring and Price 1989). The different species in this group are widely distributed across the Mojave, Sonoran, and Chihuahuan deserts of North America, and up to 10 species have been collected from a single creosote bush (Waring and Price 1989).

We sampled galls from across the current range of creosote bush and recorded the position of each sample using GPS. As many galls as possible were collected at each site and kept at room temperature until adults emerged. Following emergence, adults were preserved whole in 20 % dimethyl sulphoxide in a saturated solution of NaCl.

5.3.2 Collection of DNA Data

Genomic DNA was isolated using standard phenol chloroform methods (Hillis, Moritz et al. 1996) from single adult midges. We amplified a 451-bp fragment of the mitochondrial gene cytochrome oxidase I (*COI*) using primers C1-J-1718 and C1-N-2191 (Simon et al. 1994). We amplified a 555 bp portion of the *wingless* gene using primers 5' wg1 and 3' wg2 (Ober 2002). Resulting PCR products were purified using Shrimp Alkaline Phosphatase (SAP) and Exonuclease (EXO), and purified PCR products were used in sequencing reactions with an ABI Prism™ Dye Terminator Cycle sequencing kit.

5.3.3 Phylogenetic Analyses

Sequences for each gene for each species pair and outgroups (closely related species in the *A. auripila* group) were aligned using Clustal (Thompson et al. 1994) and adjusted by eye using Se-Al (Rambaut 1996). The best fitting model of sequence evolution was determined using Mrmodeltest (Nylander 2004). Bayesian phylogenetic analyses were performed with MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Four chains (3 heated and one cold) were run for 10×10^6 generations, and trees were sampled every 1000 generations. Bayesian posterior probabilities were employed to assess support for recovered nodes.

5.3.4 Geographic Distribution of Ancient Alleles

We constructed haplotype networks for each species using statistical parsimony as implemented in the software package TCS (Clement et al. 2000). Interior haplotypes were inferred to be relatively older alleles (Caselloe and Templeton 1994, Kimura and Ohta 1973, Miller and Schaal 2005, Posada and Crandall 2001). We visualized the spatial distribution of older alleles by plotting the locations of old and young alleles over the sampled range of creosote bush, using positional data from the collections. Old (interior) alleles of each species were enclosed in minimum spanning polygons. To reconstruct ancestral locations for each species in a species pair we also employed a maximum likelihood framework as implemented in the program PhyloMapper (Lemmon and Lemmon 2008), using a distribution of topologies and averaging of the resulting distribution of ancestral locations. We first identified the node defining each clade, and then adjusted the tree using nonparametric rate smoothing such that the branch lengths on the gene tree were proportional to time (NPRS; Sanderson 1997). We then optimized parameter values to find the joint maximum likelihood locations (latitudes and longitudes) for the node defining the clade of interest, in this case the inferred ancestral locations for each species (Lemmon and Lemmon 2008).

5.3.5 Coalescent Gene Flow and Demographic Analyses

To ensure that our dataset was appropriate for analysis in a coalescent framework, we first tested our data to ensure that the sequences were evolving in a neutral fashion, using Tajima's D (Bazin et al. 2006, Putnam 2007, Tajima 1989). We quantified per locus recombination with the composite likelihood method (Hudson 2001) as implemented in the program LDhat 2.0 (McVean et al. 2002). To ensure that our

estimation of gene flow was as conservative as possible for all gene flow analyses between *A. foliosa* and *A. auripila,* we removed the single *A. auripila* haplotype that grouped at the base of the *A. foliosa* clade. To estimate historical rates of introgression and historical population sizes between the two species pairs, we applied the 'Isolation with Migration' model implemented in the program IMa (Hey and Neilsen 2004, Neilsen and Wakeley 2001). IMa utilizes DNA sequence data from pairs of populations to infer six population-genetic parameters (population sizes for both populations and the ancestor, the divergence time, and per gene migration rates in both directions. The starting parameters and priors for these population genetic models were set as: effective population size of each population $\theta_1 = \theta_2 = 10$; the ancestral effective populations size, $\theta_a = 10$; migration between populations $m_{12} = m_{21} = 3$; and the divergence time, T=30; priors were chosen such that the posterior probability distributions of parameter estimates would be contained in the parameter space (Knowles and Carstens 2007, Won and Hey 2005). For scaling of parameter estimates we used the mutation rate of 2.3×10^{-7} from Brower et al. 1994 and 3.4×10^{-10} from Drake et al. 1999 for mitochondrial and nuclear data respectively. We searched the parameter space for each locus separately and for both loci combined using 8 Markov chains of 10 million generations each. For each parameter we recorded the marginal density as a histogram divided into 1000 equally sized bins (Won and Hey 2005). The peaks of the distributions were taken to be parameter estimates (Nielsen and Wakeley 2001, Won and Hey 2005). To establish confidence limits we used the 90% highest posterior density (HPD) for each parameter (Kronforst et al. 2006, Won and Hey 2005). We evaluated the mixing properties of MCMC by monitoring effective sample sizes (ESS) values, trend line plots of the parameter, and swapping rates between

chains. Long, well-mixing runs were repeated at least three times with different random number starting seeds. If these independent runs yielded similar posterior probability densities we considered these analyses to have converged on stationary distributions. For coalescent methods for estimating gene flow, migration rates of 0 indicate no gene flow since divergence and support a speciation model through very strong divergent selection leading to a full break in gene flow for two taxa. Alternatively, values of migration rates above zero support a sympatric model or a period of contact and gene exchange since initial divergence of the two species. For estimates of ancestral population size using IM the peaks of the posterior probability density plots for ancestral population size (Theta A) was taken to be the size of the population for each pair at divergence time (parameter T). Coalescent genealogy samplers estimate past population parameters (size, divergence times, growth rates) utilizing molecular sequence data (Kuhner 2009). We further analyzed ancestral population sizes directly for each species pair for each locus using coalescent genealogy samplers implemented in LAMARC (Kuhner 2009).

5.4 Results

5.4.1 Dataset

The dataset of *cytochrome oxidase subunit I* sequences comprised 212 sequences (97 *A. auripila*, 69 *A. foliosa*, 23 *A. rosetta*, and 23 *A. florea*). For phylogenetic analyses we included other closely related members of the *A. auripila* group (*A. clavata*, *A. pilosa*, and *A. resinosa*) as outgroup taxa (Joy and Crespi 2007). We sequenced 454 positions for the *A. foliosa* and *A. auripila*) species pair, of which 101 were parsimony informative. Pairwise differences between ingroup and outgroup species ranged from 9.2-15.2%, while differences between the two ingroup species ranged between 11-13.1%.

Intraspecific differences ranged between 0.02-4.5% for *A. foliosa* and 0-3.8% for *A. auripila*. These data were judged to be evolving in a neutral fashion by Tajima's D and Fu and Li's D tests (Tajima's D = -0.42774, P > 0.10, Fu and Li's D* = -0.63543 P > 0.10). We sequenced 451 positions for the *A. rosetta* and *A. florea* species pair, of which 54 were parsimony informative. Pairwise differences between the two ingroup species ranged between 2.66-3.62 %. Intraspecific differences ranged between 0-0.96% for *A. rosetta* and 0-0.72 % for *A. florea.* The data were judged to be evolving in a neutral fashion by Tajima's D and Fu and Li's D tests (Tajima's D = -1.03838, P > 0.1, Fu and Li's D* = 0.84396, P > 0.10). Our dataset for wingless comprised 81 sequences (17 *A. auripila*, 25 *A. foliosa*, 18 *A. rosetta*, and 21 *A. florea*). We sequenced 527 positions for both species pairs. For *A. auripila* and *A. foliosa,* pairwise differences between ingroup and outgroup species ranged from 2.6-7.8%, while differences between the two ingroup species ranged between 0.03-4.2%. Intraspecific differences ranged between 0-3.6% for *A. foliosa* and 0-2.87% for *A. auripila*. These data were judged to be evolving in a neutral fashion by Tajima's D and Fu and Li's D tests (Tajima's D = -1.11391, P > 0.10, Fu and Li's D* = -0.87363 P > 0.10). For *A. rosetta* and *A. florea*, pairwise differences between the ingroup and outgroup species ranged between 13.4-16.8%, while differences between the two ingroup species ranged between 0.064-4.9%. Intraspecific differences ranged between 0-3.31% for *A. rosetta* and 0-2.56% for *A. florea*. These data were judged to be evolving in a neutral fashion by Tajima's D and Fu and Li's D tests (Tajima's D = -1.28623, P > 0.10, Fu and Li's D* = -1.38059 P > 0.10). A summary of sequence polymorphism for both loci for both species pairs is presented in Table 1.

5.4.2 Phylogenetics

A two-partition Bayesian phylogenetic analyses converged quickly and we discarded the first 1.5 million generations. For the *COI* data, each species formed a well-supported clade displaying a posterior probability of 1.00. *A. rosetta* and *A. florea* were reciprocally monophyletic, with no shared haplotypes; all haplotypes of *A. foliosa* were also monophyletic, and all but one haplotype of *A. auripila* also formed a single clade, with a single *A. auripila* haplotype grouping as the sister lineage to *A. foliosa*. Topologies obtained for the *wingless* gene showed each species to form a reciprocally monophyletic clade, but these clades were not statistically supported, displaying posterior probabilities below 0.70.

5.4.3 Geographic Distribution of Ancient Alleles

Statistical parsimony networks were consistent with the gene trees estimated by phylogenetic methods for each gene and species pair. For the *COI* gene, minimum spanning polygons enclosing relatively old alleles indicate extensive geographic overlap of the older haplotypes of *A. auripila* and *A. foliosa* relative to younger alleles; similarly, the minimum spanning polygons of old alleles of *A. florea* and *A. rosetta* also display geographic overlap relative to younger alleles (Figure 1). For the *wingless* gene, likelihood reconstructions of ancestral locations of *A. auripila* (33.344158, -111.393298) and *A. foliosa* (33.606894, -114.175127) showed the ancestral locations of this species pair to be separated by approximately 240 km. Similarly, the likelihood reconstructions of the ancestral locations of *A.rosetta* (32.658077, -111.640821) and *A. florea* (32.898484, -112.116177) for the *wingless* gene showed the ancestral locations of this species pair to be separated by approximately 50 km. These values accord well with

ancestral locations reconstructed from the mitochondrial data *A.florea* (33.149413, -112.396116), *A.rosetta* (32.671691, -111.981717), *A.foliosa* (33.525281, -113.08206), and *A. auripila* (32.679683, -110.454436). Pairwise comparisons of results of maximum likelihood ancestral state reconstructions can be found in Table 1.

5.4.4 Coalescent Analyses

Assuming a rate of mitochondrial evolution of 2.3 % per lineage per million years (Brower 1994), the *Asphondylia foliosa* and *Asphondylia auripila* species pair diverged approximately 5 million years ago. By contrast, *Asphondylia florea* and *Asphondylia rosetta* apparently diverged much more recently - approximately 1.4 million years before present.

Repeated runs of the isolation with migration model produced marginal posterior probability distributions with clear maxima in both species-pair comparisons, for each locus and for both loci combined (Table 2, Figure 2). The results from both loci alone or in combination suggest asymmetrical gene flow, with higher rates of historical introgression from the *A. foliosa* lineage into the *A. auripila* lineage relative to the converse (Figure 2A, B, and C). By contrast, the migration parameters revealed peaks at the lower limit of resolution in both directions between the *A. florea* and *A. rosetta* pair (Figure 2D, E, and F). These peaks at zero can be interpreted as a lack of migration, indicative of no historical introgression between this species pair.

Posterior probability density (PPD) distributions for ancestral population size for both loci illustrate dramatically different ancestral population sizes for the two species pairs (θ A). The mean of the PPDs of θ A for the temporally-separated pair *A. foliosa* and

A. auripila are 112.4302, 98.4577, and 100.58 for *COI, wingless*, and both loci together respectively. Conversely, the PPDs of θ A for *A. florea* and *A. rosetta* was an order of magnitude lower, at 15.5491, 61.7321, and 39.39 for *COI, wingless,* and both loci together respectively. Estimates of ancestral effective population size from LAMARC for both mitochondrial and nuclear loci showed an order of magnitude difference in ancestral population size between the two species pairs as well. For the pair separated temporally (*A. auripila and A. foliosa*) LAMARC estimated ancestral population size ($\theta = 2N_e\mu$ for *COI* and $4N_e\mu$ for *Wg*) of 0.047 for the *COI* locus. By contrast for the pair separated by plant part (*A. florea* and *A. rosetta*) LAMARC estimated an ancestral population size of 0.006 for the *COI* locus.

5.5 Discussion

The recent implementation of coalescent-based methods provides an analytical framework to quantify levels of gene flow during divergence events (Hey and Neilsen 2004, Neilsen and Wakely 2001). The isolation-with-migration model (IM) is becoming a standard tool in the analysis of population genetic parameters associated with divergence (Hey 2005). For instance, the IM model showed no evidence for gene flow between either subspecies of common chimpanzee (*Pan troglodytes*) and bonobos (*Pan paniscus*) but a clear signal of unidirectional gene flow within the common chimpanzee, from *Pan troglodytes troglodytes* to *P. t. verus* (Won and Hey 2005). Similarly, Kronforst et al. (2006) documented asymmetrical gene flow from the butterfly *Heliconius pachinus* into *H. cydno* (2Nm = 4.326) relative to the reverse (2Nm=0.502), and unidirectional gene flow from both *H. cydno* and *H. pachinus* into *H. melpomene* (2Nm=0.294 and 0.252 respectively, Kronforst et al. 2006). More recently, Nadachowska

and Babik detected asymmetric gene flow between subspecies of the smooth newt (*Lissotriton vulgaris kosswigi* and *Lissotriton vulgaris vulgaris* in Turkey (Nadachowska and Babik 2009). Taken together, these studies suggest that divergence with some degree of gene flow may be a common feature of the speciation process.

Our study of gene flow during divergence of two sister-species of gall midge yields strongly contrasting results for the two species pairs. Thus, *A. foliosa* and *A. auripila*, which have diverged in life-history (timing of breeding) but not ecology (the plant part utilized for gall induction), form distinct, wellsupported clades, with older alleles overlapping geographically, and the gene trees suggest a history of gene flow during their divergence. The divergence of these two species therefore appears consistent with speciation in the absence of full geographical separation, or at least some period of geographical and temporal overlap involving gene exchange during divergence or upon re-contact. By contrast, *A. rosetta* and *A. florea*, which have diverged ecologically in plant-part use, but not in life-history timing, also each form distinct well-supported clades, and ancient allele mapping also shows substantial overlap of older haplotypes. However, the gene trees indicate a lack of gene flow during their divergence. These results suggest either that for this pair, differentiation and divergence were completed in allopatry and that there has been extensive geographic mixing subsequently, or that divergence took place in sympatry but the strength of divergent selection imposed by the conditions on the new plant part was relatively strong, leading to a rapid reduction in gene flow.

In this study we have compared one species pair diverged in life history timing with one species pair diverged in plant part use. The generality of inferences regarding

divergence population genetics in different ecological contexts would be greatly enhanced through further comparisons among other species pairs diverged in life history timing and other pairs diverged in plant-part use.

5.5.1 Hypotheses of Divergence

There are at least two possible scenarios that could explain the inferred pattern of divergence with gene flow in the *Asphondylia auripila* and *Asphondylia foliosa* species pair. First, an ancestral population may have diverged into the *A. foliosa* and *A. auripila* lineages through a period of spatial separation of the populations followed by a period of gene exchange. By this hypothesis, divergence starts after spatial separation and is completed during a period of gene flow following re-contact. Under the second scenario, an ancestral population diverged into *A. foliosa* and *A. auripila* through divergence of populations in the temporal dimension, at least partially in sympatry. According to this hypothesis, the lineages may have diverged with some degree of gene flow, as two lineages evolved life cycles geared to breeding after spring vs. summer rain, presumably with recurrent gene flow between populations via the same processes, such as delay or acceleration of developmental timing, that initially led to their differences in life-history. This hypothesis is consistent with the asymmetry in inferred historical gene flow, from the spring-breeding *A. foliosa* lineage into the lineage of the summer-breeding *A. auripila*.

Changes in life history timing has been shown to be important in mediating reductions in gene flowleading to speciation or host race formation, in several insect taxa, including *Rhagoletis* flies (Feder and Filchak 1999), *Eurosta* flies (Craig et al. 1993), *Enchenopa* treehoppers (Wood et al. 1990), and *Magicicada* cicadas (Wood et al. 1990;

Cooley, Simon et al. 2003). In *Rhagoletis,* differences in host-plant fruiting phenology exert divergent selection pressures on diapause and eclosion times of apple and hawthorn host races, such that hybrids emerge too early to effectively exploit one host and too late to exploit the other (Feder and Filchak 1999). In host races of *Eurosta* flies on *Solidago altissima* and *S. gigantean,* reproductive isolation is maintained through selection for synchronization with host plant phenology, leading to non-overlapping adult emergence times and preference for mating on the host plant (Craig, Itami et al. 1993; Craig, Horner et al. 2001). *Enchenopa* treehoppers eggs are rehydrated at different times on different sympatric host plants, leading to separation of adult emergence times on different host plants (Wood, Olmstead et al. 1990). In *Magicicada* periodical cicadas, large populations are thought to shift in time in response to changes in environmental cues (Cooley, Simon et al. 2003). In each of these cases, the processes of divergence necessarily take place in some degree of sympatry, which allows for gene flow during the processes of life-history specialization that lead to host-race formation or speciation.

There are several scenarios that would yield our observed pattern of a lack of gene flow in the *Asphondylia florea* and *Asphondylia rosetta* species pair combined with a small ancestral population size. In the first scenario, a novel plant part is colonized in sympatry through ovipositional mistakes (Larsson and Ekbom 1995), with eggs deposited in the new tissue type subject to different plant developmental schedules, leading to shifts in midge emergence times, thus few individuals are expected to contribute to the ancestral population. Strong selection imposed by the divergent ecological conditions on the novel host-plant tissue, such as differences in optimal depth of oviposition selecting

for different ovipositor lengths (Joy and Crespi 2007) may then lead to the rapid evolution of reproductive isolation.

By a second scenario, geographically-isolated populations evolve ecological divergence in plant-part use via some combination of drift and selection. Upon re-contact, we are left with two reproductively isolated species using different plant parts. Alternatively, reproductive isolation may evolve allopatrically under genetic drift, and upon recontact, competition between the two ecologically similar populations (using the same plant part at the same time) leads to divergence in plant-part use. These alternative hypotheses can be evaluated further by analyzing the processes involved, such as ovipositional mistakes and ecological competition, in extant populations, and by inferring the extent of historical gene flow in other pairs of *Asphondylia* sister-taxa that use the same or different plant parts.

Contrasting models of divergence within an adaptive radiation, typified by our two species-pairs, make different predictions about the influence of gene flow during divergence. Thus, divergence associated with changes in life-history timing, with retention of use of the same plant part, may involve relatively few loci, with a resultant lower scope for disruptive selection at multiple loci. For example, Prevost (1990) inferred that changes in diapause in a spruce cone midge were determined by a single dominant mutation. By contrast, divergence associated with changes in host-plant part use probably involves many loci, such that the opportunity for disruptive selection may be relatively high since the colonizing population adapts, via changes in multiple traits, to the new host-plant tissue. Pairs of sister taxa that have undergone recent ecological speciation may therefore exhibit signatures of recent disruptive selection at more loci

than species-pairs that have shifted in life-history timing. Whether taxonomic groups subject to ecological speciation should also show less historical gene flow than species pairs that have shifted due to allochronic life-history changes, or other non-ecological processes, is an open question, since patterns of gene flow will depend on the specific selective regimes and histories of allopatry and sympatry. From our example, we expect that ecological speciation, many loci, small ancestral population size, and strong selection will be associated with lower levels of historical gene flow. Additional comparative studies at the interface of population genetics and phylogenetics should help to uncover generalities linking the suites of traits that diverge with historical gene flow during adaptive radiation.

5.6 Acknowledgments

We thank Nancy Moran, Felix Breden, Christine Parent, Jeremy Bono, Tanja Schwander and Patrik Nosil for advice and comments that greatly improved this manuscript. We are also very grateful to N. Moran for laboratory space, advice and support in Arizona. Thanks to the SFU Fab* lab and Moran-Ochman group for advice and discussion. This research was funded by the Natural Sciences and Engineering Research Council of Canada and by the Society for Systematic Biology.

5.7 Literature Cited

Abrahamson, W. G., M. D. Eubanks, C. P. Blair, and A. V. Whipple. 2001. Gall flies, inquilines, and goldenrods: A model for host-race formation and sympatric speciation. Am. Zool. 41:928-938.

Becerra, J. X., and D. L. Venable. 1999. Macroevolution of insect-plant associations: The relevance of host biogeography to host affiliation. Proc. Natl. Acad. Sci. U. S. A. 96:12626-12631.

Berlocher, S. H. 2000. Radiation and divergence in the *Rhagoletis pomonella* species group: Inferences from allozymes. Evolution 54:543-557.

Caillaud, M. C., and S. Via. 2000. Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. Am. Nat. 156:606-621.

Condon, M. A., and G. J. Steck. 1997. Evolution of host use in fruit flies of the genus *Blepharoneura* (Diptera: Tephritidae): cryptic species on sexually dimorphic host plants. Biological Journal of the Linnean Society 60:443-466.

Cook, R. J., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gallwasps. Evolution 56:1821-1830.

Cooley, J. R., C. Simon, and D. C. Marshall. 2003. Temporal separation and speciation in periodical cicadas. Bioscience 53:151-157.

Craig, T. P., J. D. Horner, and J. K. Itami. 2001. Genetics, experience, and host-plant preference in Eurosta solidaginis: Implications for host shifts and speciation. Evolution 55:773-782.

Craig, T. P., J. K. Itami, W. G. Abrahamson, and J. D. Horner. 1993. Behavioral Evidence for Host-Race Formation in *Eurosta solidaginis*. Evolution 47:1696-1710.

Crespi, B. J., D. A. Carmean, and T. W. Chapman. 1997. Ecology and evolution of galling thrips and their allies. Annu. Rev. Entomol. 42:51-71.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. 2002. Speciation in the globeflower fly *Chiastocheta* spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258-68.

Dres, M., and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 357:471-492.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and plants: a study in coevolution. Evolution 18:586-608.

Emelianov, I., M. Dres, W. Baltensweiler, and J. Mallet. 2001. Host-induced assortative mating in host races of the larch budmoth. Evolution 55:2002-10.

Farrell, B. D., and C. Mitter. 1994. Adaptive Radiation in Insects and Plants - Time and Opportunity. Am. Zool. 34:57-69.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja. 2003. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. U. S. A. 100:10314-10319.

Feder, J. L., and K. E. Filchak. 1999. It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. Entomol. Exp. Et Appl. 91:211-225.

Feder, J. L., S. B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. Proc. Natl. Acad. Sci. U S A 91:7990-4.

Ferdy, J. B., L. Després, and B. Godelle. 2002. Evolution of mutualism between globeflowers and their pollinating flies. J. Theor. Biol. 217:219-234.

Funk, D. J., K. E. Filchak, and J. L. Feder. 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116:251-267.

Futuyma, D. J., M. C. Keese, and D. J. Funk. 1995. Genetic Constraints on Macroevolution - the Evolution of Host Affiliation in the Leaf Beetle Genus *Ophraella*. Evolution 49:797-809.

Gagne, R. J. 1989. The plant-feeding gall midges of North America. Cornell University Press, Ithaca.

Gagne, R. J., and G. L. Waring. 1990. The Asphondylia (Cecidomyiidae: Diptera) of creosote bush (*Larrea tridentata*) in North America. Proc. Entomol. Soc. Wash. 92:649-671.

Gagné, R. J., and P.E. Boldt. 1991. *Asphondylia* (Diptera: Cecidomyiidae) does not reflect the disjunct distribution of *Larrea* (Zygophyllaceae). Proc. Entomol. Soc. Wash. 93:509-510.

Groman, J. D., and O. Pellmyr. 2000. Rapid evolution and specialization following host colonization in a yucca moth. J. Evol. Biol. 13:223-236.

Harris, D. J., and K. A. Crandall. 2000. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda : Cambaridae): Implications for phylogenetic and microsatellite studies. Mol. Biol. Evol. 17:284-291.

Hawkins, B. A., R. D. Goeden, and R. J. Gagne. 1986. Ecology and taxonomy of the *Asphondylia* spp. (Diptera: Cecidomyiidae) forming galls on *Atriplex* spp. (Chenopodiaceae) in Southern California. Entomography 4:55-107.

Hillis, D. M., C. Moritz, and B. K. Mable. 1996. Molecular Systematics (2nd Edition). Sinauer and Associates

Hunter, K. L., J. L. Betancourt, B. R. Riddle, T. R. Van Devender, K. L. Cole, and W. G. Spaulding. 2001. Ploidy race distributions since the Last Glacial Maximum in the North American desert shrub, Larrea tridentata. Glob. Ecol. Biogeog. 10:521-533.

Jaenike, J. 1989. Genetics of Butterfly Hostplant Associations. Trends Ecol. Evol. 4:34-34.

Jaenike, J. 1990. Host Specialization in Phytophagous Insects. Annu. Rev. Ecol. Syst. 21:243-273.

Jermy, T. 1984. Evolution of insect/host plant relationships. Am. Nat. 124:609-630.

Jones, R. G., R. J. Gagne, and W. F. Barr. 1983. Biology and taxonomy of the *Rhopalomyia* gall midges (Diptera: Cecidomyiidae) of *Artemesia tridentata* Nutall (Compositae) in Idaho. Contrib. Am. Entomol. Inst. 21:1-76.

Mardulyn, P., M. C. Milinkovitch, and J. M. Pasteels. 1997. Phylogenetic analyses of DNA and allozyme data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. Syst. Biol. 46:722-747.

Nyman, T., A. Widmer, and H. Roininen. 2000. Evolution of gall morphology and host-plant relationships in willow-feeding sawflies (Hymenoptera : Tenthredinidae). Evolution 54:526-533.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-8.

Ronquist, F., and J. Liljeblad. 2001. Evolution of the gall wasp-host plant association. Evolution 55:2503-22.

Thompson, J. N. 1994. The Coevolutionary Process. University of Chicago Press.

Waring, G. L., and P. W. Price. 1989. Parasitoid pressure and the radiation of a gallforming group (Cecidomyiidae: *Asphondylia* spp.) on creosote bush (*Larrea tridentata*). Oecologia 79:293-299.

Wood, T. K., K. L. Olmstead, and S. I. Guttman. 1990. Insect Phenology Mediated by Host-Plant Water Relations. Evolution 44:629-636.

Table 11 Summary of nucleotide variation.

Species	Locus	N[a]	Sites[b]	Hd[c]	Pi[d]
A. auripila	COI	97	454	0.896	0.01577
	Wingless	17	527	0.709	0.00531
A. foliosa	COI	69	454	0.934	0.02482
	Wingless	25	527	0.641	0.01785
A. rosetta	*COI*	23	451	0.802	0.00332
	Wingless	18	527	0.928	0.01146
A. florea	COI	23	451	0.628	0.00184
	Wingless	21	527	0.7	0.06287

[a] Total number of sequences.

[b] Seqence length (bp).

[c] Haplotype diversity.

[d] Nucleotide diversity per site.

Table 12 Pairwise comparison of maximum-likelihood estimates (MLE) for reconstructions of ancestral locations of each species in each pair.

Comparison	***Wingless***	***COI***	**Distance**
A.auripila-A.foliosa	242 km	264 km	NA
A.florea-A.rosetta	52 km	65 km	NA
A.auripila-A.auripila	NA	NA	115 km
A.foliosa-A.foliosa	NA	NA	102 km
A.rosetta-A.rosetta	NA	NA	31 km
A.florea-A.florea	NA	NA	38 km

Table 13 Maximum-likelihood estimates (MLE) and 90% highest posterior density (HPD) intervals of IMa model parameters; effective population size of species 1 (θ_1), effective population size of species 2 (θ_2), effective size of ancestral population (θ_a), migration rate of species 1 into species 2 (m_{12}) and migration rate of species 2 into species 1 (m_{21}). HPD90Lo is the lower bound of the estimated 90% highest probability density (HPD) interval and HPD90Hi is the upper bound of the estimated HPD interval.

Comparison	**q1**	**q2**	**qa**	**m1**	**m2**
A.foliosa-A.auripila (COI)	82.0981	65.4911	na	0.778	0.302
HPD90Lo	65.1166	55.1274	na	0.578	0.206
HPD90Hi	100.4531	76.1046	na	0.978	0.398
A.florea-A.rosetta (COI)	3.7884	2.2731	na	0.09	0.122
HPD90Lo	1.413	0.5939	na	0.001	0.001
HPD90Hi	6.6554	4.3618	na	0.311	0.455
A.foliosa-A.auripila (Wg)	3.6975	2.1966	83.8428	2.756	2.068
HPD90Lo	2.0986	0.0999	48.4669	1.204	0.428
HPD90Hi	199.7637	4.2971	162.1894	6.804	7.124
A.florea-A.rosetta (Wg)	34.2439	5.6624	31.188	0.001	0.001
HPD90Lo	16.6276	2.4267	8.3588	0.001	0.001
HPD90Hi	61.3874	11.9539	110.6412	0.105	0.201
A.foliosa-A.auripila (COI+Wg)	29.1443	24.7605	100.5844	0.734	0.602
HPD90Lo	20.539	18.9154	50.9011	0.522	0.418
HPD90Hi	37.9119	31.4174	160.8214	1.09	0.838
A.florea-A.rosetta (COI + Wg)	25.2676	7.1233	39.3933	0.0365	0.0705
HPD90Lo	14.7327	3.3659	1.7105	0.001	0.003
HPD90Hi	39.6734	13.0773	104.0116	0.135	0.245

Figure 12 Maximum Likelihood reconstructions of ancestral locations for the wingless gene (Wg) and cytochrome oxidase subunit I (COI) and minimum spanning polygons of ancient alleles for the COI locus. A: *Asphondylia foliosa* (open polygon) and *A. auripila* (hatched polygon) and B: *A. rosetta* (open polygon) and *A. florea* (hatched polygon). Sampling locations are denoted by black circles and the range of *Larrea tridentata* is outlined in gray. Note reconstructed ancestral locations for both species pairs occur in close proximity relative to the range of the host-plant species.

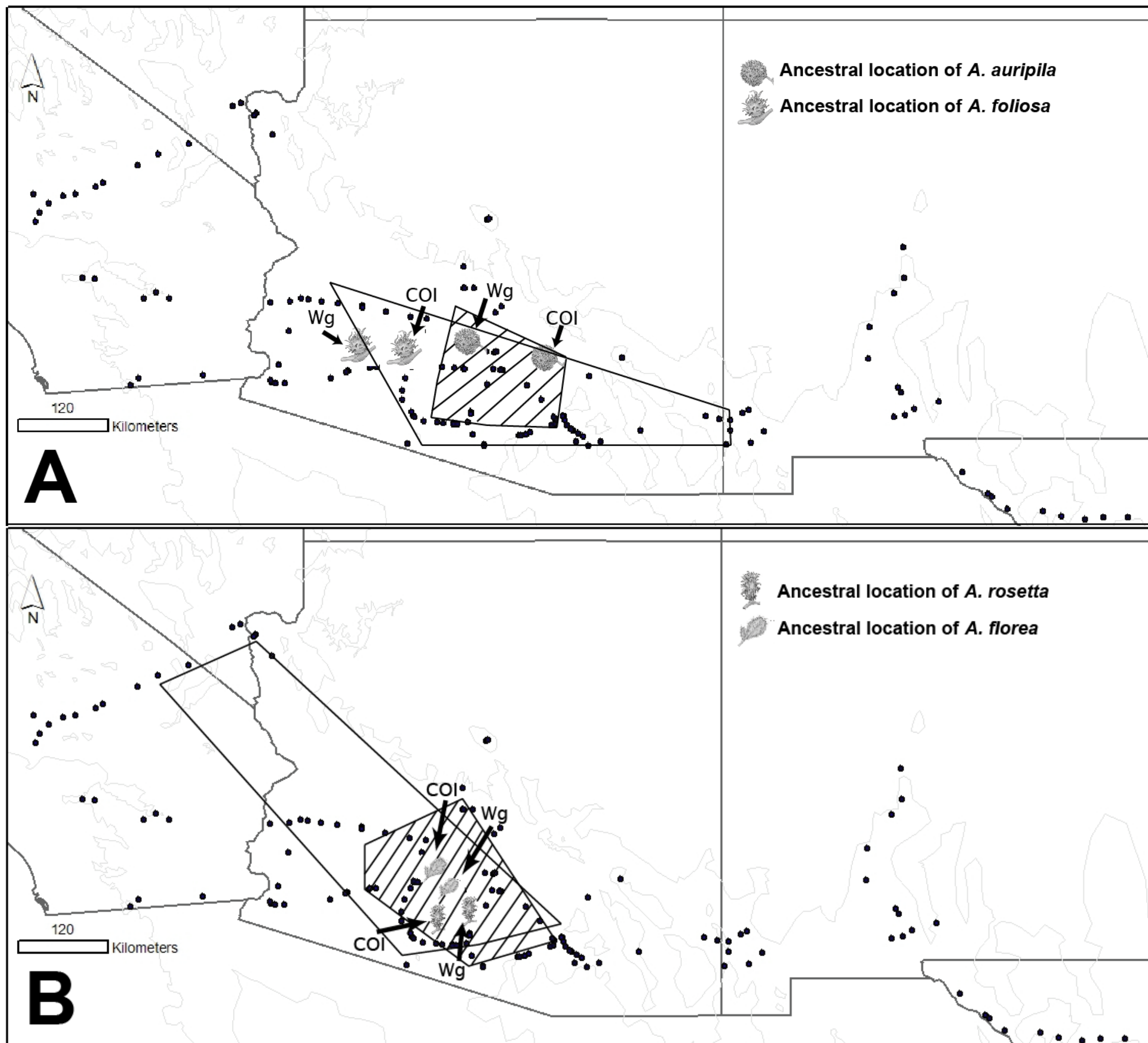


Figure 13 The marginal posterior probability distributions for between species migration rates estimated with IMa. Bidirectional migration is shown for pairwise comparisons for the cytochrome oxidase subunit I gene (A, B), wingless gene (C, D), and both loci together (E, F) for the species pair which displays a temporal shift (Asphondylia foliosa and A. auripila); and no migration is detected for pairwise comparisons for both COI and Wg for the species pair which displays a plant-part shift (A. florea and A. rosetta).

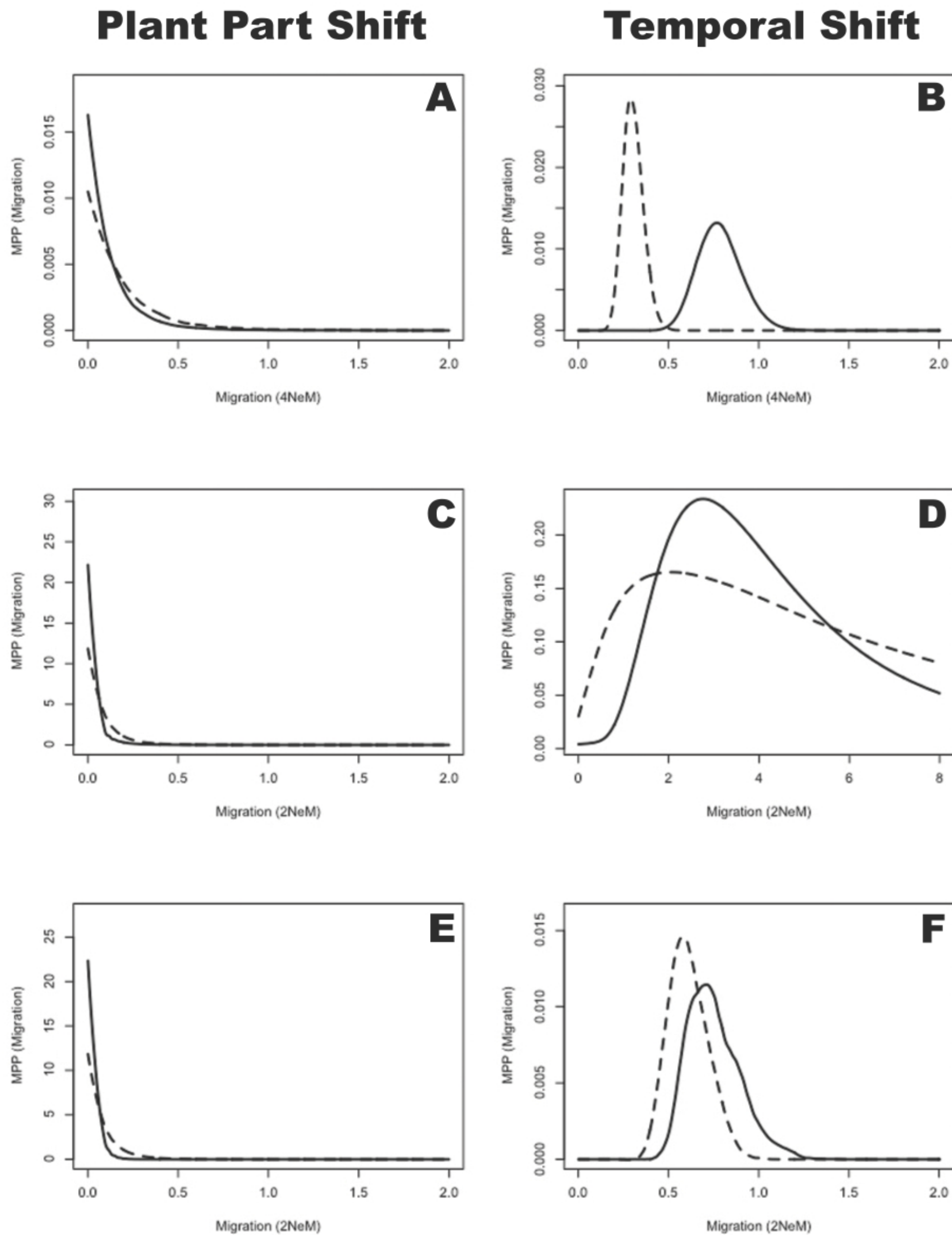
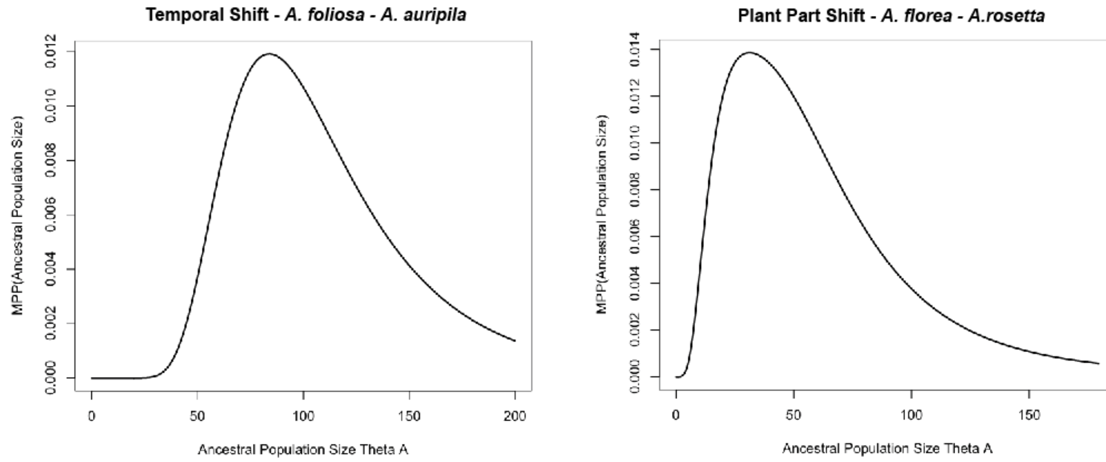


Figure 14 The marginal posterior probability distributions for ancestral population size estimated from the wingless gene for both species pairs using IMa; note larger ancestral population size for the species pair diplaying divergence in life history timing relative to the species pair diverged in plant-part use.



Chapter 6: Major Findings and Conclusions

6.1 Summary

Previous macroevolutionary models of phytophagous insect diversification have predominantly considered single processes in isolation (e.g. plant species richness, Novotny et al. 2006). However, the causes of diversification among phytophagous insects are multitudinous, such that a macroevolutionary framework which allows the simultaneous consideration of a variety of potential factors should be preferred over one which considers one factor alone. Such combined macro and microevolutionary analyses reveal that diversification of cecidomyiid flies may be explicable in part from the effects of island biogeographic processes.

Examination of cecidomyiid fly adaptive radiations within differing island biogeographic contexts reveals host shifts to occur among plants which are closely related and within-plant speciation to be an important process in insular contexts. While results of analyses within this framework have proven useful in explaining the diversification of phytophagous cecidomoyiid flies, the generality of the usefulness of this framework can be tested through comparative analyses among other groups of cecidomyiid flies and among other phytophagous insect groups.

Speciation without a host-plant shift in both *Asphondylia* and *Rhopalomyia* is associated with shifts among plant parts and shifts among time periods, indicating that these two types of shifts may be general mechanisms facilitating divergence and adaptive radiation within a single host-plant species. Focused examination of gene flow,

population demographic history, and selection on insect morphology reveals the importance of evolutionary processes involved in within host-plant divergence.

6.2 Nexus of Macro and Micro Evolutionary Processes

The likelihood of speciation by specialized phytophagous insects during adaptive radiation as a result of shifts to new niches may be driven by a non-linear relationship between the opportunity for gene flow and the strength of divergent natural selection (Figure 15). Thus, whereas divergence is unlikely or rare between distantly-related host-plant species because such shifts require substantial adaptation, divergence is much more probable among relatively closely-related host-plants because the balance between the strength of natural selection and gene flow favours this process. Thus, shifts among related host plants in diverse plant families may provide optimum conditions for divergence, where selection is strong but the homogenizing effects of gene flow are reduced. Shifts among resources on the same host plant likely require substantially less adaptation because many of the selective forces will be relatively similar to those in the originating population: for example, host-plant chemistry and defences are expected to be more similar within a host plant relative to between host plants. Thus, selection against hybrids may be reduced between populations on the same host plant species because the strength of divergent selection is expected to be lower and the opportunity for gene flow is expected to be higher, relative to between-host shifts.

Comparison of two radiations of cecidomyiid flies reveal that *Rhopalomyia* which inhabit a non-insular host-plant species are nearly twice as diverse as the *Asphondylia* which inhabit an insular host-plant species over roughly the same time frame, suggesting the possibility that adaptive radiation of phytophagous insects resulting from a

combination of host shifting and within-plant speciation may be more rapid than within-plant speciation alone. Future studies comparing speciation rates among phytophagous insects which have radiated predominantly through host-plant shifts with those which have speciated within hosts will provide further insight into the pace of adaptive radiation in phytophagous insects in these different contexts.

6.3 Within Host Speciation: Plant-part Shifts

Divergence in plant-part use appears to be a common mechanism or outcome of speciation in the absence of a shift to a novel host-plant species (Cook et al. 2002; Joy and Crespi 2007; Condon et al. 2008). Divergence in plant-part use may result in part from competition during or after speciation. Thus, Cook et al. (2002) showed that speciation of *Andricus* gall wasps is more commonly associated with shifts to a novel part of the same host plant than with shifts between different host plant species. *Andricus* gall wasps are thought to have undergone within host-plant divergence through intraspecific competition for oviposition sites, such that oviposition on the 'wrong' organ of the correct oak species more often result in gall formation than oviposition on the wrong oak species (Cook et al. 2002).

In contrast to this hypothesis, divergence in plant part use may evolve, in part, in association with selection for better control, by the insect, over nutrient availability (Nyman et al. 2000). Previous work on willow-feeding sawflies showed galling to evolve from positions on leaf tissue to more central locations of the host plant such as stems, or buds (Nyman et al. 2000). The observation that radiations of narrowly host-specific phytophagous insects begin through colonization of leaf tissues and subsequently progress to more central plant parts may be a consequence of the ease of colonizing leaf

tissue relative to other plant tissues (Nyman et al. 2000). Alternatively, the greater abundance of phytophages inhabiting leaf tissue relative to other tissue types (Bernays and Chapman 1994) may make observing leaf-to-leaf shifts more likely relative to stem-to-stem shifts. Shifts to more-central locations on the plant apparently involve selection for better control over plant nutrient flow, evasion of abscission, and larger gall and brood size (e. g., Inbar et al. 2004).

Colonization sequence may also dictated in part by the degree of similarity in host-plant traits (such as chemistry, morphology) between the source plant and the novel plant. If colonizing leaf tissue is easier relative to more central parts of the host plant (such as the stem), then colonization of more insular host plants may often commence with the leaf. The constraint of beginning colonization of a host plant with the leaf tissue may be relaxed among related plants with more similar host plant characteristics (i.e. plant chemistry, plant defenses). This idea could be tested by comparing the sequence of plant parts colonized among a sample of closely-related plant species relative to a sample of distantly-related plant species. By this hypothesis I expect the colonization of distantly related plants to consistently begin with the leaf and progress towards more central parts of the plant; by contrast, colonization of closely related plants would be expected to occur at peripheral or central parts of the host plant with similar frequency.

6.4 Within Host Speciation: The Importance of Time

Temporal shifts, the colonization of new time periods of plant growth or other alterations in life history timing which facilitate reproductive isolation, have been shown to be generally important in mediating reductions in gene flow leading to speciation or host race formation in many insect taxa, including *Rhagoletis* flies (Feder and Filchak

1999), *Eurosta* flies (Craig et al. 1993), *Enchenopa* treehoppers (Wood et al. 1990), and *Magicicada* cicadas (Cooley et al. 2003). Similarly, divergence in life-history timing is a consistent theme among studies implicating divergence within a single host-plant species in speciation. Phenological divergence may be facilitated by shifts to competition-free space, in that the insects that have shifted in time are likely to be released from competition (e.g. for oviposition sites; Craig et al. 2000). In *Chiastocheta* flies inhabiting *Trollius* species, Després et al. (2002) demonstrated that diversification has involved both shifting hosts and radiation within a host, and that within-host diversification may be a result of competition for oviposition or feedings sites favoring temporal shifts in oviposition timing and shifts to different larval food resources (Ferdy et al. 2002). The prolonged diapause of the gall midge *Dasineura rachiphaga* is thought to be a mechanism that evolved in the context of selection for reduced intraspecific competition for limiting oviposition sites (Prévost 1990). Conversely, temporal shifts concomitant to within-host divergence may be driven by natural enemies, such that insects displaying altered life-history timing may escape the community of parasitoids present in the ancestral population. Along these lines Clancy and Price (1986) suggested that natural enemies may have driven temporal shifts in a willow-feeding *Pontania* sawfly to maximize survival and enemy free space.

That divergence in life-history timing in both *Artemisia*- and *Larrea*- associated cecidomyiid flies is connected with speciation within a single host-plant accords well with previous studies of other phytophagous insects. As with other phytophagous insects, by facilitating a reduction or break in gene flow, divergence in life-history timing in cecidomyiid flies may be an important part of the divergence process. Alternatively,

divergence in life-history timing during within host-plant speciation of cecidomyiid flies may be a result of post speciational competition for the same resource or selection imposed by natural enemies. The observation that in both *Rhopalomyia* and *Asphondylia* within host-plant speciation events associated with retention of the same part of the plant also involve divergence in life-history timing suggests that the reduction in gene flow mediated by changes in life-history timing may be important in these instances. The hypothesis that divergence in timing may be more important in within plant speciation when the same part of the plant is retained could be tested further by comparing the role of life-history timing between cases of plant-part divergence and plant-part retention in replicated studies of other groups of phytophagous insects.

6.5 Within Host Speciation - Generality?

Within host-plant speciation may be most frequent when divergence in plant-part use (increasing the strength of natural selection) coincides with temporal shifts (reducing the effects of gene flow). Consistent with this idea, Condon and Steck (1997) found that specialization on different host plant tissues concomitant with divergence in life-history timing has shaped the evolution of host-plant use in *Blepharoneura* flies. Within-host speciation in these flies has been associated with specializing on either male or female flowers. Female and male flowers are temporally isolated, such that shifts from one sex-flower to another could result in isolation of plant part-specific populations.

Phytophagous insect diversity estimates have generally been based upon counts of morphologically distinguishable insect species collected from plants (Novotny et al. 2006). However, molecular evidence suggests that counts of morphospecies dramatically underestimate both diversity and host specificity (Condon et al. 2008; Monaghan et al.

2009). Evidence from *Blephanoneura* fruit flies (Condon et al. 2008) and *Strobilomyia* flies (Sachet et al. 2006) shows that some of this cryptic diversity arises through within host-plant speciation. As molecular data accumulate from additional phytophagous insects, a more precise understanding of the relative frequency of within- versus between-host shifts will emerge.

6.7 Final Remarks

Macroevolutionary analyses presented in this thesis show that consideration of the diversification of phytophagous insects within the framework of island biogeographic theory can have great explanatory power.

Further, combined phylogenetic, population genetic, ecological and morphological analyses such as those employed here provide a powerful, integrated approach to understanding how phytophagous insects, the most speciose of animal groups, have evolved.

As outlined above, future studies of rates of diversification, colonization sequence, the relative importance of host-plant shifts versus within plant shifts, and the importance of divergence in life-history timing, will further advance our understanding of macro- and micro-evolutionary processes involved in phytophagous insect diversification.

6.8 Literature Cited

Bernays, E.A., and Chapman, R.F. 1994. *Host plant selection by phytophagous insects.* Chapman and Hall.

Clancy, K. and P.W. Price. 1986. Temporal variation in three-trophic-level interactions among willows, sawflies, and parasitoids. Ecology 67:1601-1607.

Condon, M. A., and G. J. Steck. 1997. Evolution of host use in fruit flies of the genus *Blepharoneura* (Diptera: Tephritidae): cryptic species on sexually dimorphic host plants. Biol. J. Linn. Soc. 60:443-466.

Condon, M.A., S.J. Scheffer, M.L. Lewis, and S.M. Swensen. 2008. Hidden neotropical diversity: greater than the sum of its parts. Science 320:928-931.

Cook, R. J., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gallwasps. Evolution 56:1821-1830.

Cooley, J. R., C. Simon, and D. C. Marshall. 2003. Temporal separation and speciation in periodical cicadas. Bioscience 53:151-157.

Craig, T. P., J. K. Itami, W. G. Abrahamson, and J. D. Horner. 1993. Behavioral evidence for host-race formation in *Eurosta solidaginis*. Evolution 47:1696-1710.

Craig, T. P., J. K. Itami, C. Shantz, W. G. Abrahamson, J. D. Horner, and J. V. Craig. 2000. The influence of host plant variation and intraspecific competition on oviposition preference and offspring performance in the host races of *Eurosta solidaginis*. Ecol. Entomol. 25:7-18.

Després, L., E. Pettex, V. Plaisance, and F. Pompanon. (2002) Speciation in the globeflower fly Chiastocheta spp. (Diptera: Anthomyiidae) in relation to host plant species, biogeography, and morphology. Mol. Phylogenet. Evol. 22:258–268.

Feder, J. L. and K.E. Filchak. 1999. It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. *Entomol. Exp. Appl*. 91:211–225.

Ferdy, J. B., L. Després, and B. Godelle. 2002. Evolution of mutualism between globeflowers and their pollinating flies. J. Theor. Biol. 217:219-234.

Funk, V. A., R.J. Bayer, S. Keeley, R. Chan, L. Watson, B. Gemeinholzer, E. Schilling, J.L. Panero, B.G. Baldwin, N.T. Garcia-Jacas, A. Susanna, and R.K. Jansen. 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. In: Friis, I. and H. Balslev [eds.], Proceedings of a Symposium on Plant Diversity and Complexity Patterns - Local, Regional and Global Dimensions. The Royal Danish Academy of Sciences and Letters, Copenhagen. *Biol. Skr*. 55:343-374.

Inbar, M., M. Wink, and D. Wool. 2004. The evolution of host plant manipulation by insects: molecular and ecological evidence from gall-forming aphids on *Pistacia*. Mol. Phylogenet. Evol. 32:504-511.

Monaghan, M.T., R. Wild, M. Elliot, T.Fujisawa, M. Balke, D.J.G. Inward, D.C. Lees, R. Ranaivosolo, P. Eggleton, T.G. Barraclough, and A.P. Vogler. 2009. Accelerated species inventoryon Madagascar using coalescent based models of species delineation. Syst. Biol. 58:298-311.

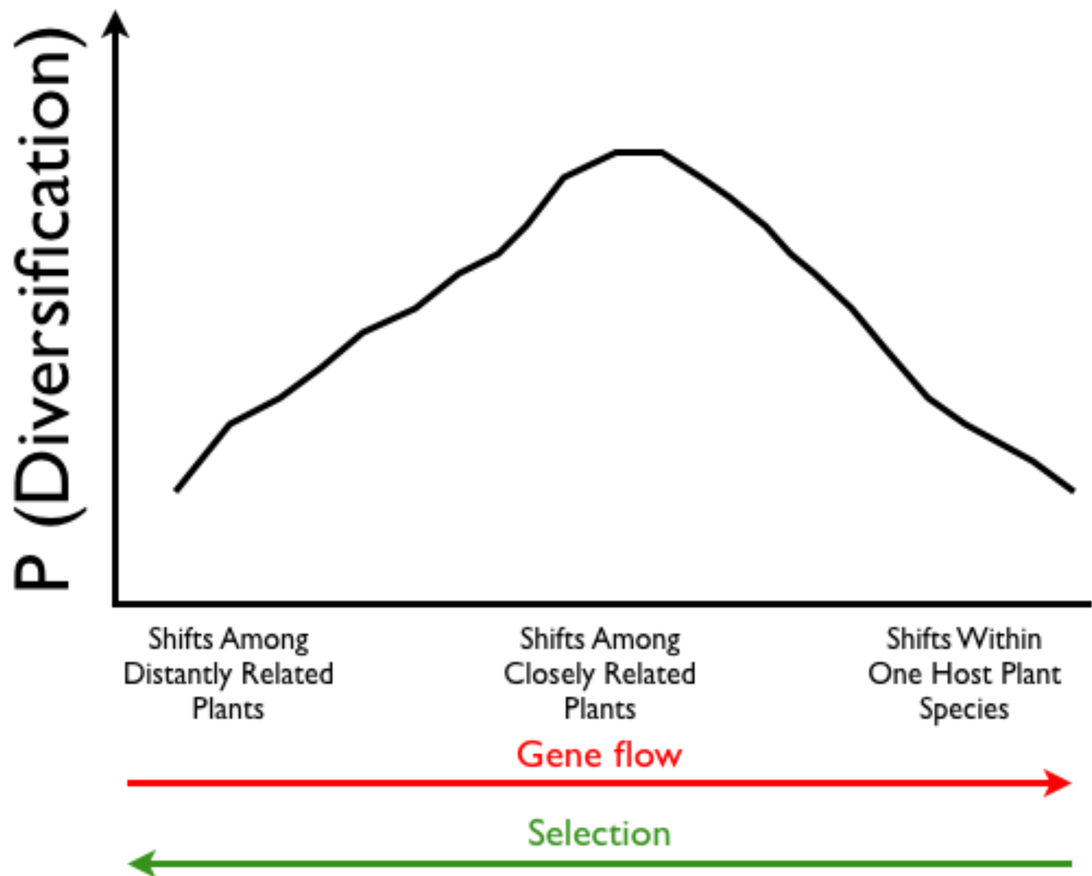
Novotny, V., P. Drozd, S.E. Miller, M. Kulfan, M. Janda, Y. Basset, and G.D. Weiblen. 2006. Why are there so many species of herbivorous insects in tropical rainforests? *Science* 313:1115-1118.

Nyman, T., A. Widmer, and H. Roininen. 2000. Evolution of gall morphology and host-plant relationships in willow-feeding sawflies (Hymenoptera : Tenthredinidae). Evolution 54:526-533.

Prevost, Y. H. 1990. Spruce cone axis midge, *Dasineura rachiphaga* Tripp (Diptera, Cecidomyiidae), in cones of black spruce, *Picea mariana* (Mill) Bsp. Can. Entomol. 122:441-447.

Wood, T. K., K. L. Olmstead, and S. I. Guttman. 1990. Insect phenology mediated by host-plant water relations. Evolution 44:629-636.

Figure 15 Predicted probability of the divergence of a lineage of specialized phytophagous insects in different contexts. Shifts among distantly related host plant species are unlikely because substantial adaptation is required and there is little new variation coming in to the population for selection to act on because gene flow is low. Shifts among closely related plants are likely because some adaptation is required and gene flow is not high enough to swamp the effects of divergent selection. Shifts within a host plant species are unlikely because divergent selection is lowest and the homogenizing effects of gene flow are highest.



Appendix

Appendix 1: Comparative mitochondrial genomics reveals truncated transfer RNA genes are a common feature of Cecidomyiid mitochondrial genomes

Modified from Beckenbach and Joy (2009) by permission of Oxford University Press

S1.1 Abstract

We determined the complete mitochondrial genome sequences of two species of gall midges (Diptera: Cecidomyiidae), as well as partial sequence from a third cecidomyiid and a species from a related family, the Sciaridae. The sciarid sequence has a number of rearrangements of tRNA genes, relative to other dipterans, but is otherwise unremarkable. In contrast, the cecidomyiid genomes possess a number of very unusual features. First, the two complete sequences are very small compared with other dipteran mitochondrial genomes. The genome of *Mayetiola destructor* is only 14,759 bp while that of *Rhopalomyia pomum* is only 14,503 bp, comparable with genome sizes observed in some arachnids. Second, all three cecidomyiid species have very high A + T content - more than 83% for the coding region. Third, all three cecidomyiid species possess a number of rearrangements of tRNA genes, including variations within the family. Fourth, the most extraordinary feature of cecidomyiids examined in this study is an extreme truncation of all tRNA genes, including the loss of TWC arms and apparent absence of the 3' part of the aminoacyl stems.

The truncated tRNA genes of cecidomyiids are very similar to those previously reported for spiders and appear to represent a second, independent origin of these structural features. It is likely that they are made functional through RNA editing, perhaps using the 5' end of the aminoacyl stem as a template for the construction of the required 3' end.

S1.2 Introduction

Animal mitochondrial genomes encode about 37 essential genes, including 12–13 genes that code for components of the electron transport system, and a minimal translation system, which includes 22 transfer RNA (tRNA) and two ribosomal RNA (rRNA) genes (Wolstenholme 1992; Boore 1999). The genomes are very compact, usually without introns, and with few noncoding residues outside of a single control region. Despite the apparent simplicity of animal mitochondrial genomes, their structure and function often exhibit a number of unusual molecular features, including alternative genetic codes, RNA editing, and a diverse array of tRNA structures (Burger et al. 2003).

Typically, tRNAs fold into a cloverleaf secondary structure, with an aminoacyl (or acceptor) stem, a DHU stem and loop, an anticodon stem and loop, a TC stem and loop, and a smaller variable loop separating the anticodon stem from the TC stem. Although the majority of mitochondrial tRNAs have this standard structure, variations on the basic structure have been found throughout the animal kingdom. Examples have been encountered missing the DHU stem, or the TC stem, or both (Masta and Boore 2008). Nematodes, such as *Caenorhabditis elegans*, have what appears to be the minimal functional tRNA structure, lacking both the DHU and TC stems (Wolstenholme 1992). But the most extraordinary modifications of tRNA genes have been found in spiders

(Masta 2000; Masta and Boore 2004, 2008; Qiu et al. 2005). Not only are the TC stem and loop missing from all tRNA genes but also the 3' portion of the aminoacyl stem evidently is not coded in the DNA. Because tRNAs cannot function without a paired aminoacyl stem, it is evident that some form of RNA editing must exist to construct the 3' end de novo. Lavrov et al. (2000) showed that mismatches in the aminoacyl stem of a centipede are corrected by RNA editing using the 5' end as a template. Masta and Boore (2004) noted that this mechanism could be used to reconstruct the entire 3' end of the aminoacyl stem of spider tRNAs, using the 5' end as template. They provided a number of predictions concerning the evolution of this editing capability. These predictions include the possible relaxation of constraints on the sequence of the aminoacyl stems and a general evolutionary trend toward the loss of the 3' end from all tRNA genes in the genome.

During preliminary studies of the mitochondrial genome of a gall midge (the Hessian fly, *Mayetiola destructor*), we encountered tRNA-like structures similar to those of spiders. We therefore undertook the sequencing of the complete genome of this species, along with a more extensive examination of flies from this interesting family.

Cecidomyiid flies (Arthropoda: Diptera: Nematocera: Bibionomorpha: Sciaridae: Cecidomyiidae) are an ancient lineage of flies known to exist for more than 150 My (Yukawa and Rohfritsch 2005). Cecidomyiidae underwent explosive diversification in the Cretaceous period coincident with the appearance of angiosperm plants and are now a hyperdiverse family encompassing more than 5,700 described species. The predominant life history mode, as the name gall midge implies, involves the induction of gall structures on various plant tissues followed by larval feeding and development within the

galls. They are known to diversify both through host plant shifts (Price 2005) and through ecological partitioning of a single plant (Joy and Crespi 2007; Stireman et al. 2008).

Cecidomyiid flies are known to have some unusual genetic features. Quantification of the size of the nuclear genome of *M. destructor* showed it to be the smallest known nuclear genome of any insect at 0.09 pg (Gregory et al. 2007). The most unusual features of cecidomyiids involve the behavior of chromosomes during meiosis and early development (White 1973). Although the details vary widely across the family, features include the absence of homologous pairing of chromosomes and the formation of a highly asymmetrical spindle during spermatogenesis, as well as the elimination of chromosomes from somatic tissues (but not the germ line) during early development. In *M. destructor*, for example, the germ line carries about 40 chromosomes, but as a result of chromosome elimination during cleavage, the somatic tissues of both sexes have only about eight chromosomes. Sex determination occurs by the differential elimination of X chromosomes from somatic tissues (White 1949). Some of these features, including unipolar spindle formation during spermatogenesis, the elimination of chromosomes during early cleavage, and sex determination by elimination of X chromosomes, are shared by the Dipteran family Sciaridae. These genetic features, together with morphological synapomorphies, support a sister relationship between the two families (White 1949; Wood and Borkent 1989; Oosterbroek and Courtney 1995).

In this study we examine the evolutionary extent of truncated tRNA genes within the family Cecidomyiidae by identifying tRNA genes in three species in different genera. We also compare the number and type of gene rearrangements among these cecidomyiid mitochondrial genomes. To place the evolution of the truncation in tRNA genes into a

broader context within the Cecidomyiidae, we infer phylogenetic relationships of the taxa under study and among the family more generally using mitochondrial cox1 sequences. All family- or subfamily-level phylogenies for Cecidomyiidae to date have been based exclusively on morphological characters (Gagné 1989; Roskam 2005). No single previous phylogeny of any sort has encompassed all the genera under study here. We also include partial sequencing of the mitochondrial genome of *Bradysia amoena*, of the related family Sciaridae.

S1.3 Methods

S1.3.1 Source and Collection of Specimens

Adults of the Hessian fly *M. destructor* (Cecidomyiidae) were obtained from a laboratory culture of R. Shukle, Purdue University. Adults of B. amoena (Sciaridae) were obtained from a laboratory culture maintained by S. Gerbi, Brown University. Both species were preserved in 95% EtOH and provided through the Dipteran Tree of Life Project. We collected specimens of the stem galling *Asphondylia rosetta* (Cecidomyiidae) from *Larrea tridentata* (creosote bush) near Tucson, AZ, and specimens of the leaf galling *Rhopalomyia pomum* (Cecidomyiidae) were collected from *Artemesia tridentata* (sagebrush) near Kamloops, BC, Canada. We obtained the sequences of the tRNALeu(UUR) gene from GenBank for *Asteromyia carbonifera* and *A. euthamiae* (accession numbers EU439835 and EU439782) for comparison with the homologous sequences obtained in this study. No other identified tRNA sequences were available for representatives of this family.

S1.3.2 Genome Sequencing

Individual specimens were ground in the presence of protease K, and total genomic DNA was extracted using a standard phenol–chloroform extraction protocol. After ethanol precipitation, extracts were dried and dissolved in 100–200 µl of distilled water. The general strategy for amplification and sequencing was to amplify fragments of 500–1,500 bp using standard primers (Simon et al. 2006). Details of the amplification conditions and purification of templates are given in Beckenbach and Stewart (2009). Initial attempts with standard primer pairs yielded only a few fragments, scattered about the genomes. In particular, primers based in tRNA genes invariably failed to amplify. Additional sequence was obtained by primer walking using taxon-specific primers based on preliminary sequence, paired with standard primers, or with other taxon-specific primers. Two regions proved most challenging: the region between the nad3 and nad5 genes and the control region, between the small ribosomal subunit and the nad2 gene. We were successful in amplifying across these regions in *Mayetiola* and *Rhopalomyia* and completed these sequences using primer walks. Primer sequences and locations are available from the authors. Repeated attempts to amplify across these regions in *Asphondylia* and *Bradysia* were unsuccessful.

S1.3.3 Annotation of the Sequences

Sequences were assembled manually, based on regions of overlap and on the locations of amplification and sequencing primers. Protein-coding genes were identified as open reading frames, and by alignment with homologous sequences of other Diptera. The rRNA genes were identified by alignment with sequences of other arthropods. Identification of tRNA genes posed the greatest challenges. Gene junctions having

unassigned sequence were scanned online using tRNAscan-SE (Lowe and Eddy 1997), with a cove score cutoff of 1. This process found tRNA genes for *Bradysia* but generally failed for cecidomyiid sequences. Where putative tRNA genes were located in the cecidomyiids, the sequences overlapped downstream genes and showed mismatches in the aminoacyl stem. Examination of these regions suggested that cecidomyiids possessed truncated tRNA sequences, similar to those observed in some arachnids (Masta 2000; Masta and Boore 2008). We used an approach described in Masta and Boore (2004) for locating the genes. We used the following criteria: 1) a well-formed and well-paired anticodon stem and loop, with an appropriate anticodon; 2) a well-formed and well-paired DHU stem and loop, with 3–4 bp in the stem; and 3) at least nine residues upstream of the DHU stem that cannot be assigned to an upstream gene coded on the same strand. This last criterion assures that the 5' end of the aminoacyl stem is present after processing of the primary transcripts. The most crucial criterion, however, was the conservation of the putative DHU stems and of the entire anticodon stem–loop sequence across the three cecidomyiid species examined here.

S1.3.4 Phylogeny Reconstruction and Character Evolution

To characterize the evolution of tRNA truncation within Cecidomyiidae, we inferred a phylogenetic tree and delineated instances of tRNA truncation at the tips. The tree was inferred using combined mitochondrial sequence data from our sequenced genomes, and a 444-bp fragment of the cox1 gene obtained from GenBank for species in all available genera; adequate coverage was not available for any other genes. Sequences were aligned using Clustal (Thompson et al. 1994) and adjusted by eye using Se-Al (Rambaut 1996). The best fitting model of sequence evolution was determined using

Modeltest (Posada and Crandall 1998). We also employed MrModeltest 2.2 (Nylander 2004) to determine the best model for use in Bayesian phylogeny estimation. We reconstructed phylogenetic relationships among cecidomyiid species under maximum likelihood (ML) and maximum parsimony (MP) using PAUP* 4.0b10 (Swofford 2002). We also reconstructed phylogenetic relationships using Bayesian methods as implemented in MrBayes 3.12 (Ronquist and Huelsenbeck 2003). Two parallel runs utilizing default priors, four heated chains, while sampling trees from one cold chain every 1,000 generation were run for 10 million generations.

To test the hypothesis that all species displaying the tRNA truncation were monophyletic, we employed the Shimodaira–Hasegawa (SH) test and the Templeton test as implemented in PAUP* (Swofford 2002) to compare the best tree with a constraint tree which forces monophyly of cecidomyiid species known to have truncated tRNA genes.

S1.3.5 Data Deposition

Sequences have been deposited in GenBank under the following accession numbers: *Mayetiola destructor*, GQ387648 [GenBank] ; *Rhopalomyia pomum*, GQ387649 [GenBank] ; *Asphondylia rosetta*, GQ387650 [GenBank] ; and *Bradysia amoena*, GQ387651.

S1.4 Results

We determined complete mitochondrial genome sequences for two species of gall midges, *M. destructor* and *R. pomum*, as well as partial sequences for a third cecidomyiid, *A. rosetta*, and a sciarid, *B. amoena*. In all four species, the protein-coding

and rRNA genes are in the typical arthropod positions and orientation, but rearrangements involving tRNA genes are evident in all four genomes. The cecidomyiid sequences exhibit a number of unusual features. The most interesting is a severe truncation of all tRNA gene sequences, which is described in detail below.

S1.4.1 tRNA Gene Rearrangements

The two complete cecidomyiid sequences show a number of rearrangements involving tRNA genes (Figure 16). The tRNA Ile gene has been inverted and transposed in both *Mayetiola* and *Rhopalomyia* from the typical arthropod position between the control region and the nad2 gene to the block of tRNA genes between nad3 and nad5. The tRNA Asn gene has been moved from the nad3–nad5 block to a position between the tRNA Gly gene and nad3 in both species. The tRNA Glu gene, located within the nad3–nad5 region, has been inverted in both species, relative to the typical arthropod gene arrangement. As these rearrangements involve regions of the genome not determined in *Asphondylia*, we cannot say whether they are shared by this species.

Both the tRNA Thr and tRNA Pro genes, located between the nad4l and nad6 genes, are inverted in all three cecidomyiid species. This observation is particularly interesting, as it requires a minimum of two separate steps because they remain in the same position relative to the typical arthropod genome arrangement. Either they both underwent inversions independently or there was a single inversion involving both genes, followed by a transposition. The two genes are transcribed from different strands, both in the typical gene arrangement and in the inverted arrangement found in cecidomyiids. In *Rhopalomyia*, the genes are separated by a noncoding block consisting of a tandem repeat

of 11 copies of an 18-bp sequence. The genes overlap by 16 bp in *Asphondylia* and 26 bp in *Mayetiola*.

The tRNATyr gene appears in three different places in the three cecidomyiid species examined here. In *Rhopalomyia* it is retained in the typical arthropod position, between the tRNACys gene and cox1. In *Mayetiola* it has been transposed to the nad3–nad5 tRNA block, whereas in *Asphondylia* it has been transposed to a position between tRNA Ser(UCN) and nad1. It is coded on the minority strand in all three genomes. This gene is also moved from the typical arthropod position in the sciarid *Bradysia*, although it is evidently not located within the regions sequenced in this study.

A number of other rearrangements are evident in the partial *Asphondylia* sequence. Genes not present in their typical positions include tRNACys, tRNA Leu(UUR), tRNA Lys, and tRNA Asp. Rearrangements evident in the partial *Bradysia* sequence include tRNA Cys, tRNA Tyr, and tRNA Leu(UUR), which are not present in their typical positions. We assume they have been moved to positions within the regions not sequenced in this study. In addition, tRNA Arg and tRNA Asn, usually located within the nad3–nad5 sequence block, are identifiable between the nad6 and cytb genes in *Bradysia*. In the typical arthropod mitochondrial gene arrangement the nad6 and cytb genes abut.

S1.4.2 Features of Cecidomyiid Mitochondrial Genomes

The two complete cecidomyiid sequences show a number of unusual features. Both genomes are smaller than most insect mitochondrial genomes and comparable to those observed in some spiders (Masta and Boore 2004, 2008). The genome of *Mayetiola*

is 14,759 bp, including a control region of about 600 bp, whereas the genome of *Rhopalomyia* is only 14,503 bp, with a control region of about 360 bp. Notably, they represent the smallest known dipteran mitochondrial genomes and are comparable to those found among many arachnids (Masta and Boore 2008).

Part of the reduction in overall genome size can be attributed to a reduction in length of most of the protein-coding genes (Table 14). Nearly all the protein-coding genes are shorter than those of other dipterans, especially the NADH dehydrogenase complex genes. This reduction cannot be ascribed to differing views of the annotation of these genes, as all protein-coding genes in *Mayetiola* and all but two in *Rhopalomyia* have DNA-encoded terminators, and in most cases the start codons are unambiguous.

Both the tRNA Thr and tRNA Pro genes, located between the nad4l and nad6 genes, are inverted in all three cecidomyiid species. This observation is particularly interesting, as it requires a minimum of two separate steps because they remain in the same position relative to the typical arthropod genome arrangement. Either they both underwent inversions independently or there was a single inversion involving both genes, followed by a transposition. The two genes are transcribed from different strands, both in the typical gene arrangement and in the inverted arrangement found in cecidomyiids. In *Rhopalomyia*, the genes are separated by a noncoding block consisting of a tandem repeat of 11 copies of an 18-bp sequence. The genes overlap by 16 bp in Asphondylia and 26 bp in *Mayetiola*.

The tRNATyr gene appears in three different places in the three cecidomyiid species examined here. In *Rhopalomyia* it is retained in the typical arthropod position, between the tRNACys gene and cox1. In *Mayetiola* it has been transposed to the nad3–

nad5 tRNA block, whereas in *Asphondylia* it has been transposed to a position between tRNASer(UCN) and nad1. It is coded on the minority strand in all three genomes. This gene is also moved from the typical arthropod position in the sciarid *Bradysia*, although it is evidently not located within the regions sequenced in this study.

A number of other rearrangements are evident in the partial *Asphondylia* sequence. Genes not present in their typical positions include tRNACys, tRNALeu(UUR), tRNALys, and tRNAAsp. Rearrangements evident in the partial *Bradysia* sequence include tRNACys, tRNATyr, and tRNALeu(UUR), which are not present in their typical positions. We assume they have been moved to positions within the regions not sequenced in this study. In addition, tRNAArg and tRNAAsn, usually located within the nad3–nad5 sequence block, are identifiable between the nad6 and cytb genes in *Bradysia*. In the typical arthropod mitochondrial gene arrangement the nad6 and cytb genes abut.

S1.4.3 Features of Cecidomyiid Mitochondrial Genomes

The two complete cecidomyiid sequences show a number of unusual features. Both genomes are smaller than most insect mitochondrial genomes and comparable to those observed in some spiders (Masta and Boore 2004, 2008). The genome of *Mayetiola* is 14,759 bp, including a control region of about 600 bp, whereas the genome of *Rhopalomyia* is only 14,503 bp, with a control region of about 360 bp. Notably, they represent the smallest known dipteran mitochondrial genomes and are comparable to those found among many arachnids (Masta and Boore 2008).

Part of the reduction in overall genome size can be attributed to a reduction in length of most of the protein-coding genes (Table 14). Nearly all the protein-coding genes are shorter than those of other dipterans, especially the NADH dehydrogenase complex genes. This reduction cannot be ascribed to differing views of the annotation of these genes, as all protein-coding genes in *Mayetiola* and all but two in *Rhopalomyia* have DNA-encoded terminators, and in most cases the start codons are unambiguous.

The three cecidomyiid genomes examined in this study have extremely high A + T contents, ranging from just over 74% in cox1 to over 90% in the atp8 and nad6 genes (Table 14). These values are considerably higher than those that have been observed in other dipterans and are comparable to those of the honeybee, *Apis mellifera* (Crozier RH and Crozier YC 1993). Overall A + T content of *Mayetiola* is 84.1%, including 83.6% for the coding region and 90.5% for the control region. The values for *Rhopalomyia* are 85.2% overall, 84.6% coding and 94.2% for the control region. Regions sequenced from *Asphondylia* show comparable A + T content, whereas those in the sciarid *Bradysia* have A + T content more typical of other dipterans (Table 14).

S1.4.4 Structural Characteristics of Cecidomyiid tRNAs

The tRNAs coded in all three species evidently lack TC stem–loop structures, as well as the 3' end the aminoacyl (acceptor) stem. Structures similar to these have been previously observed in spiders (Masta and Boore 2004, 2008; Qiu et al. 2005). Evidence that some arthropods have evolved a mechanism to reconstruct the 3' end of the aminoacyl stem through RNA editing, presumably using the 5' end as a template, was provided by Lavrov et al. (2000).

The sequences of the tRNAs identified in this study are given in Figure 17 and Figure 19, and 20 ; examples of our interpretation of the folding structures are given in Figure 17. Folded structures of all tRNAs are given in Figure 18-20. The regions including the entire anticodon stem and loops are very well conserved across the cecidomyiid species in all 22 tRNAs. In nearly all instances where a nucleotide substitution is observed in the aminoacyl stem, a compensatory substitution retains pairing capability. The DHU stems are also well conserved across each species, again with compensatory substitutions observed in some cases. Most of the variations observable in the DHU stem–loop regions are within the loops, including both nucleotide substitutions and indels (Figure 17 and 19).

In contrast to the conservation of the anticodon stem–loops and DHU stems, there is little conservation either upstream or downstream from these structures. We assume that nine residues are required upstream of the DHU stem, including two unpaired residues and seven residues needed to form the 5' end of the aminoacyl stem. In no case do these nine residues overlap an upstream gene coded on the same strand, and in many instances this putative 5' stem region follows immediately after the terminator codon of an upstream protein-coding gene. We take these observations as evidence of a functional role for these nine residues.

S1.4.5 Phylogenetic Results

Figure 18 illustrates the reconstructed phylogenetic relationships among genera in the subfamily Cecidomyiinae (family Cecidomyiidae). All available data for Cecidomyiidae are from this subfamily. The MP, ML, and Bayesian analyses yielded trees of very similar topology. The support for the recovered nodes was generally robust

toward the tips and declined with depth in the tree, apparently due to mutational saturation of the cox1 locus at this level of divergence.

SH test (difference in –ln L = 79.95058, $P < 0.001$) and Templeton test (difference in length = 47, $P < 0.001$) strongly reject the hypothesis that taxa sharing the character state of truncated tRNA genes form a monophyletic group relative to other genera.

S1.5 Discussion

Complete sequences of the mitochondrial genomes from two representatives of gall midges (family Cecidomyiidae), as well as partial sequence from a third, show that the genomes are highly modified in several ways. They have a very high A + T content, a general reduction of overall length compared with most other animal mitochondrial genomes, rearrangement of some of the tRNA genes, and, most notably, truncation of all tRNA genes. The overall length reduction is partly due to a shortening of most of the major genes and to a severe truncation of the tRNA genes. Phylogenetic tests support the placement of four genera displaying truncated tRNA genes in disparate parts of the cecidomyiid phylogeny, indicating either repeated evolution within the family or, more likely, origin of the truncation mechanism in a common ancestor shared by all cecidomyiine species. The presence of fully coded tRNA genes in *Bradysia* (family Sciaridae) implies that truncated tRNA structures arose after the separation of these two families. Further investigation of the structure of tRNA genes in the Cecidomyiidae will provide insights into the evolutionary origins of truncated tRNA genes.

S1.5.1 Rearrangement of tRNA Genes

All the taxa examined in this study have tRNA rearrangements relative to the typical arthropod mitochondrial genome organization. These rearrangements include inversions and transpositions. All three cecidomyiid genomes share inversions of both the tRNAThr and tRNAPro genes. These changes appear to simplify transcription and processing of genes in this region, by bringing together the N-strand–encoded genes from tRNAThr to nad5 and the J-strand genes from tRNAPro to tRNASer(UCN) into continuous, uninterrupted blocks. Mapping these changes onto the phylogeny of the subfamily Cecidomyiinae (Figure 18) shows that these changes occurred early in the diversification of this subfamily. The inversion and transposition of tRNAIle gene, the inversion of the tRNAGlu gene, and the transposition of the tRNAAsn gene in both *Mayetiola* and *Rhopalomyia* indicate that these rearrangements occurred prior to the separation of these two genera. These genes were not located within the regions sequenced in *Asphondylia*, so we cannot pinpoint the origin of these changes on this phylogeny.

Several tRNA gene rearrangements have occurred since the separation of *Mayetiola* and *Rhoplalomyia*. These changes include transposition of tRNATyr in *Mayetiola* from its typical position, which is retained in *Rhopalomyia*, and transposition of a small block, which includes the tRNAGlu and tRNAPhe genes. These two genera are placed in the same Tribe by morphology and appear as sisters in our cox1 molecular phylogeny. The number of changes that we observe in tRNA gene organization, however, suggests that they are not close on an absolute timescale.

The observation of multiple shared, as well as taxon-specific, tRNA gene rearrangements indicates that tRNA gene arrangements may provide useful markers for more detailed phylogenetic reconstruction in the Cecidomyiidae. The sciarid examined in this study, *Bradysia*, also possesses tRNA gene rearrangements, suggesting that gene arrangements may be generally useful for phylogenetic reconstruction in this section of the superfamily Sciaridae.

S1.5.2 Truncated tRNA Genes

Sequences interpreted here as tRNA genes are well conserved in the anticodon stems and loops, and in the DHU stems. This conservation is evident even where the tRNA genes appear in different places in the genome, as we observe for the tRNATyr gene. Despite strong conservation of two of the arms of the standard cloverleaf structure, there is little or no evidence of TC stems or loops, or sequence corresponding to the 3' end of the aminoacyl stem. Nonetheless, the presence of well-formed anticodon stems and loops corresponding to all 22 expected tRNA genes in both *Mayetiola* and *Rhopalomyia* strongly suggests a coding role for these sequences. As they occupy regions between the protein-coding and rRNA genes and no other coding role is evident for these sequences, we conclude they are the functional tRNA genes. Further, the sequence conservation evident in the anticodon loops and DHU stems would degrade rapidly in the mutation-prone (Lynch et al. 2006) mitochondrial genome. Intergenic residues where no coding role is evident are not conserved among these species. Thus, selection likely maintains the conserved tRNA sequences in cecidomyiid mitochondrial genomes, further supporting a coding role for them. Certainly, the conservation would not persist over the

millions of years of evolution, which separates the cecidomyiid genera displaying truncated tRNA genes if they were not functional.

The truncation of these genes poses problems in locating and annotating the tRNA genes (Masta and Boore 2004). In the absence of a well-paired acceptor stem, the 3' end is not clearly defined. The region downstream from the anticodon stem is extremely variable in sequence and length. In some cases possible stem structures can be found downstream, but such structures are not consistent and often overlap downstream genes. For example, the region downstream of the anticodon stem in tRNAAsp in *Rhopalomyia* could be folded in several ways, but the region is absent in Mayetiola (Figure 19). Similarly, a region that could be folded in tRNAMet from *Mayetiola* is missing from *Rhopalomyia*. Rather than hypothesize a variety of structural differences among the tRNA genes and between the same tRNA gene in different cecidomyiid species, it is more reasonable to assume that all the tRNA genes function in a similar manner.

The absence of well-paired aminoacyl stems poses interesting questions regarding the processing of primary transcripts in cecidomyiids. In all Metazoa where transcription has been documented, the primary transcripts are polycistronic and must be processed to yield the necessary messenger (mRNA), tRNA, and rRNA transcripts. The tRNA punctuation model was proposed to account for the processing in the human mitochondrial genome (Ojala et al. 1981). In this model, the tRNA sequences are removed from the primary transcripts, producing the mature mRNA and rRNA transcripts in the process. There is evidence supporting this model in *Drosophila melanogaster* (Stewart and Beckenbach 2009). It is generally assumed that the secondary structure of the tRNA sequences provides the required signals for processing of the primary

transcripts (Ojala et al. 1981; Clary and Wolstenholme 1985). In the absence of standard cloverleaf tRNA structures, some modification of this model appears necessary. It is possible, of course, that the anticodon stem alone provides the appropriate signals.

A second implication of the tRNA punctuation model is that genes coded on the same strand cannot overlap because transcripts for adjacent genes are derived from the same processing events. The only exceptions are the atp8/atp6 and nad4l/nad4 genes, which are translated from bicistronic transcripts (Berthier et al. 1986). In the cecidomyiid sequences examined here, 11 of the 22 tRNA genes have an adjacent downstream gene coded on the same strand (Figure 16). If the tRNA punctuation model holds for cecidomyiids, the processed transcripts for these tRNAs must be severely truncated at the 3' end prior to any editing steps.

The 5' ends of the aminoacyl stems are poorly conserved (Figure 17). Most are extremely A + T rich, so it is often possible to find an A + T–rich sequence downstream that will pair as many as five of the seven residues we have assigned to the 5' end of the aminoacyl stem. In some cases potential matching regions are within downstream genes. Rather than postulate different structures for some of the tRNA genes and for the same gene in different species, it seems more reasonable to hypothesize that all genes function in a common manner.

The most likely mechanism for proper functioning of truncated tRNA genes is RNA editing, using the 5' end of the aminoacyl stem as a template for the construction of a well-paired stem (Lavrov et al. 2000; Masta and Boore 2004). Editing of the 3' end of the aminoacyl stem has been demonstrated in snail mitochondrial tRNA sequences, as well (Yokobori and Pääbo 1995). The tRNA genes in snails, where the most extensive

editing is required, are those that overlap with a downstream gene. It is likely that they are actually truncated at the start of the downstream gene and reconstructed using the 5' end as a template. If this mechanism is used, it assures a fully matched stem regardless of the 5' end sequence and would appear to relax the constraints on the actual sequence of those stems (Masta and Boore 2004). Our observations support this prediction (Figure 16).

There is, however, another potential source of constraint on the aminoacyl stem sequence: the requirement for proper recognition of the tRNA by amino acid–charging enzymes. In both prokaryotes and nuclear-encoded eukaryote systems, tRNA-synthetase recognition is based on determinants (specific residues) in both the aminoacyl stem and anticodon loop (Giege et al. 1998; Beuning and Musier-Forsyth 1999). The lack of conservation of the 5' end of the tRNAs in cecidomyiids would seem to require either that proper recognition of the tRNA during the charging process relies solely on determinants in the anticodon stem and loop or a rapid co-evolution of aminoacyl determinants with the aminoacyl synthase genes.

S1.5.3 Parallel Evolution of Truncated tRNA Genes

The finding of truncated tRNA genes in cecidomyiids implies that this feature has arisen at least twice within the Arthropoda. Masta and Boore (2008) provided evidence of multiple losses of TC arms within arachnids, but loss of both the TC arm and 3' portion of the aminoacyl stem from all tRNA genes may have a single origin in spiders. A second independent origin within the cecidomyiids supports the prediction that development of mechanisms to edit the 3' end of tRNAs and the ability of the ribosome to accommodate

tRNAs lacking the TC arms can lead to the pattern of truncation of tRNA genes observed in spiders and cecidomyiids (Masta and Boore 2004).

The observation that all, not just a subset, the tRNA genes in both spiders and gall midges have become truncated suggests that natural selection may favor truncation, once a mechanism evolves which allows the proper functioning of these genes. Several hypotheses might favor the development of such a mechanism. One hypothesis is that the mechanism evolves as a result of selection for a smaller genome size (Dufresne et al. 2005; Giovannoni et al. 2005, 2008). Selection may favor smaller genomes to enhance replication speed in organisms, which must develop rapidly to take advantage of ephemeral conditions. The small size of both the nuclear and mitochondrial genomes of cecidomyiids is consistent with this idea. A second hypothesis is that the RNA-editing mechanism may evolve as a way to ameliorate the mutational consequences of asexuality in asexual organellar genomes (Lynch 1997; Lynch et al. 2006). A third hypothesis is that marked population subdivision can result in extremely small local effective population size, allowing a variety of deleterious mutations to be fixed. A final hypothesis states that modification of the ribosome in a common ancestor may result in the relaxation of rules governing interactions between ribosome and tRNA molecules.

S1.5.4 Phylogenetic Origin of Truncated tRNA Genes in Cecidomyiids

Our phylogeny and results of SH and Templeton tests for the cox1 gene support the hypothesis that taxa displaying reduced genome size and truncated tRNA genes do not form a compact monophyletic group relative to other available genera in the family. Monophyly of the taxa under study here would also be unsupported by phylogenies inferred from morphological characters (Roskam 2005). Inferences about cecidomyiid

phylogeny based on morphological characters place *Asphondylia* and *Mayetiola* species in highly divergent clades (Roskam 2005). Further, *Mayetiola* and *Asphondylia* are morphologically very different and display dramatic differences in life history. Thus, multiple lines of evidence (genetic, morphological, and life history) support the premise that cecidomyiid taxa, which have evolved tRNA truncation, are widely dispersed across the subfamily. These results imply either that the mechanism which allows truncation of the tRNA genes evolved once in a common ancestor, or that it has evolved multiple times, convergently, within the family. We favor a single origin in the Cecidomyiidae.

The truncated and rearranged tRNA genes shown in this study illustrate the dynamic nature of cecidomyiid mitochondrial genomes and extend the taxonomic breadth of the observation that in some lineages tRNA genes are severely truncated to Diptera. As more data become available on both the distribution of truncated tRNA genes and the mechanisms that allow them to function, it will become more tractable to test hypotheses about the roles of various evolutionary forces favoring the development of the mechanism.

S1.6 Acknowledgements

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada to ATB, and National Science Foundation (US) grant EF-0334949, Building the Dipteran Tree of Life, Brian Wiegmann, PI. We thank the SFU FAB* lab for helpful discussion and comments.

S1.7 Literature Cited

Beckenbach AT, Stewart JB. 2009. Insect mitochondrial genomics 3: the complete mitochondrial genome sequences of representatives from two neuropteroid orders: a dobsonfly (order Megaloptera), a giant lacewing and an owlfly (order Neuroptera). Genome 52:31-38.

Berthier F, Renaud M, Alziari S, Durand R. 1986. RNA mapping on *Drosophila* mitochondrial DNA:precursors and template strands. Nuc. Acids Res. 14:4519-4533.

Beuning PJ, Musier-Forsyth K. 1999. Transfer RNA recognition by aminoacyl-tRNA synthetases. Biopolymer 52:1-28.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27:1767-1780.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. Trends Genetics 19:709-716.

Cameron SL, Dowton M, Castro LR, Ruberu K, Whiting MF, Austin AD, Diement K, Stevens J. 2008. Mitochondrial genome organization and phylogeny of two vespid wasps. Genome 51:800-808.

Clary DO, Wolstenholme DR. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22:252-271.

Dufresne A, Garczarek L, Partensky F. 20005. Accelerated evolution associated with genome reduction in a free-living prokaryote. Genome Biol. 6: R14.

Frazer-Abel AA, Hagerman PJ. 2008. Core flexibility of a truncated metazoan mitochondrial tRNA. Nuc. Acids Res. 36:5472-5481.

Geige R, Sissler M, Florentz C. 1998. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res. 26:5017-5035.

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC, Mathur EJ. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science 3091242-1245.

Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho JC, Oh HM, Kitner JB, Vergin KL, Rappe MS. 2008. The small genome of an abundant coastal ocean methylotroph. Env. Microbiol. 10:1771-1782.

Gregory TR, Nicol JA, Tamm B, Kullman B, Kullman K, Leitch IJ, Murray BG, Kapraun DF, Greilhuber J, and Bennett MD. 2007. Eukaryotic genome size databases. Nucleic Acids Research 35:D332-D338.

Joy JB, Crespi B. 2007. Adaptive radiation of gall-inducing insects within a single host-plant species. Evolution 61:784-795.

Lavrov DV, Brown WM, Boore JL. 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. Proc. Natl. Acad. Sci. USA. 97:13738-13742.

Lowe TD, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955-964.

Lynch M. 1997. Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes. Mol. Biol. Evol. 14:914-925.

Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. Science 311:1727-1730.

Masta SE. 2000. Mitochondrial sequence evolution in spiders: Intraspecific variation in tRNAs lacking the TΨC arm. Mol. Biol. Evol. 17:1091-1100.

Masta SE, Boore JL. 2004. The complete mitocondrial genome sequence of the spider *Habronattus oregonensis* reveals rearranged and extremely truncated tRNAs. Mol Biol. Evol. 21:893-902.

Masta SE, Boore JL. 2008. Parallel evolution of truncated transfer RNA genes in arachnid mitochondrial genomes. Mol. Biol. Evol. 25:949-959.

Nylander JA. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala Univ, Sweden.

Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470-474.

Oosterbroek P, Courtney G. 1995. Phylogeny of the nematocerous families of Diptera (Insecta). Zool. J. Linnean. Soc. 115:267-311.

Posada D, Crandall KA. 1998. MODELEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Price PW. 2005. Adaptive radiation of gall-inducing insects. Basic Appl Ecol. 6:413-421.

Qiu Y, Song D, Zhou K, Sun H. 2005. The mitochondrial sequences of *Heptathela hangzhouensis* and *Ornithoctonus huwena* reveal unique gene arrangements and atypical tRNAs. J. Mol. Evol. 60:57-71.

Rambaut A. 1996. Se-Al: Sequence alignment editor. Available at http://evolve.zoo.ox.ac.uk/.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Roskam H. 2005.Phylogeny of gall midges (Cecidomyiidae). In: Raman A, Schaefer CW, Withers TM. (eds) Biology, ecology, and evolution of gall-inducing arthropods. Science Publishers. Inc., Enfield, New Hampshire, pp 205-320.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 37:545-579.

Stewart JB, Beckenbach AT. 2009. Characterization of mature mitochondrial transcripts in *Drosophila*, and the implications for the tRNA punctuation model. Submitted to Gene.

Stireman JO, Janson EM, Carr TG, Devlin H, Abbot P. 2008. Evolutionary radiation of *Asteromyia carbonifera* (Diptera: Cecidomyiidae) gall morphotypes on the goldenrod *Solidago altissima* (Asteraceae). Biol. J. Linn. Soc. Lond. 95:840-858.

Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony(*and other methods). Ver. 4. Sinauer Associates, Sunderland, MA.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc. Acids Res. 22:4673-4680.

White MJD. 1949. Cytological evidence on the phylogeny and classification of the Diptera. Evolution 3:252-261.

White MJD. 1973. *Animal Cytology and Evolution*. Third Ed. Cambridge Univ Press. New York.

Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. In Jeon KW, Wolstenholme DR.(eds). *Mitochondrial Genomes*. International Review of Cytology 141:173-216.

Wood DM, Borkent A. 1989. Phylogeny and classification of the Nematocera. In McAlpine JF (ed.) Manual of Nearctic Diptera, vol. 3. Agriculture Canada Monograph 32:1333-1370.

Yokobori S-I, Pääbo S. 1995. Transfer editing in land snail mitochondria. Proc. Natl. Acad. Sci. USA. 92:10432-10435.

Yukawa J, Rohfritsch O. 2005. Biology and ecology of gall inducing Cecidomyiidae (Diptera). In: Raman A, Schaefer CW, Withers TM. (eds) Biology, ecology, and evolution of gall-inducing arthropods. Science Publishers. Inc., Enfield, New Hampshire, pp 273-304.

Table 14 Characteristics of protein coding genes of cecidomyiids and other Diptera. The length of each gene is given in codons. Initiators and terminators are shown as coded in the DNA sequences. Overlaps of one or more residues in the terminator with the downstream gene are indicated with parentheses.

Genus	atp6	atp8	cox1	cox2	cox3	cytb	nad1	nad2	nad3	nad4	nad4l	nad5	nad6
Mayetiola													
Length	223	51	511	225	260	375	301	324	116	436	90	558	160
Initiator	ATG	ATT	ATT	ATA	ATA	ATT	ATG	ATA	ATT	ATT	TTT	ATT	ATG
Terminator	TAA	(TAA)	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TA(A)	TAA
A+T (%)	82.5	94.1	74.3	80.2	79.2	79.3	82	89.8	86	82.3	89.4	83.2	91.1
Rhopalomyia													
Length	223	51	512	224	260	372	301	325	116	437	90	554	155
Initiator	ATG	ATT	TTT	ATA	TTA	ATA	ATA	ATA	ATA	ATA	ATA	ATT	ATA
Terminator	TAA	(TAA)	TA(A)	TAA	TA(A)	T	TAA	TAA	TAA	T	TAA	TAA	TAA
A+T (%)	82.8	(TAA)	76.6	81.5	80.3	80.3	83.3	91.3	85.8	83.1	89	85.2	92.1
Asphondylia													
Length	223	52	511	219	n.a.	378	n.a.	n.a.	n.a.	437	90	n.a.	154
Initiator	ATG	ATT	ATT	ATT	ATT	ATA	ATA	n.a.	n.a.	ATG	TTT	ATA	ATT
Terminator	TAA	(TAA)	TAA	TAA	n.a.	TAA	TAG	TAA	n.a.	TAA	TAA	n.a.	TAA
A+T (%)	83	91.7	74.8	81.1	80.8	81.1	n.a.	n.a.	n.a.	85.2	88.6	84.8	92.9
Bradysia													
Length	224	55	512	227	262	378	313	344	n.a.	446	n.a.	574	n.a.
Initiator	ATG	ATC	ATG	ATT	ATG	ATA	ATA	ATA	ATT	ATG	n.a.	ATA	n.a.
Terminator	TAA	(TAA)	TAA	TAA	TAA	TAA	TAA	TAG	TAA	TAA	(TAA)	n.a.	TAA
A+T (%)	75	88.5	69	74.3	71.4	73	76.7	81.9	74.7	78.4	n.a.	78.7	n.a.
Anopheles													
Length	226	53	512	228	262	378	314	341	117	448	99	574	173
Initiator	ATG	ATC	TCG	ATG	ATG	ATG	ATA	ATC	ATA	ATG	ATG	TAT	ATT
Terminator	TA(A)	(TAA)	T	T	TA	TAA	TAA	T(AA)	TAA	T	TAA	TAA	TAA
A+T (%)	74.3	81.8	68.6	73.1	70.4	72.4	76.6	83	79.4	77.7	82.7	78.2	84.9
Drosophila													
Length	224	53	511	228	262	378	315	341	117	447	96	573	174
Initiator	ATG	ATT	TCG	ATG	ATG	ATG	ATA	ATT	ATT	ATG	ATG	ATT	ATT
Terminator	TA(A)	(TAA)	TAA	T	TAA	TAA	T	T(AA)	TAA	T	TAA	T	TAA
A+T (%)	75.8	82.4	69.6	73.8	71.1	73.8	78.6	81.4	79.4	79.6	83.8	77.7	84.8

Figure 16 Organization of the mitochondrial genomes of representatives of the families Cecidomyiidae and Sciaridae. The genome structures are linearized to place the control region at the end. The “standard” gene order is that found in Drosophila, and is widespread in insects. The upper line of gene names are coded on the majority (“J”) strand; the lower line gives those coded on the minority (“N”) strand. Protein coding genes: A6, A8 are atpase subunits 6 and 8; C1-C3 are cytochrome oxidase subunits; CB is cytochrome B; and N1-N6, N4L are NADH dehydrogenase subunits. Ribosomal genes: 16s and 12s are the large and small subunits. The tRNA genes are indicated by their single letter amino acid designations. Lines indicate transpositions; curved arrows indicate that an inversion is involved.

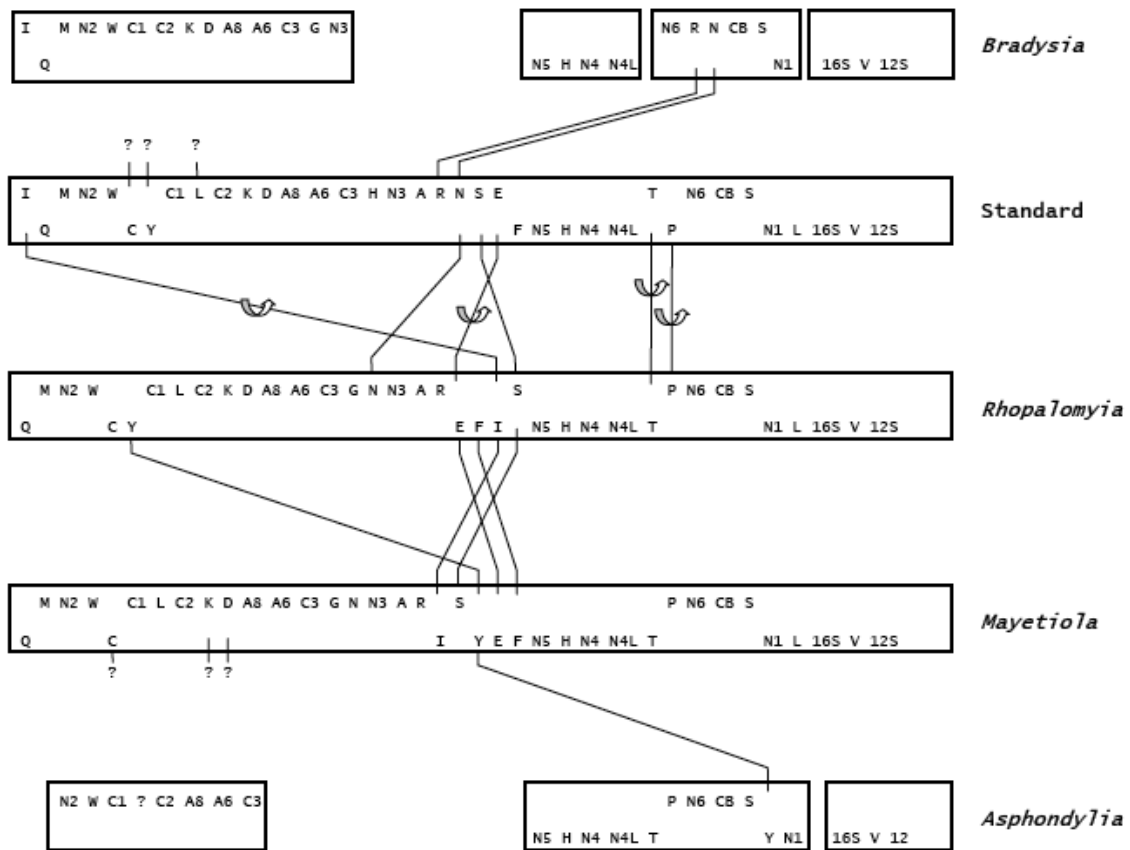


Figure 17 Examples of inferred tRNA secondary structure from each cecidomyiid genome. Note consistent lack of a TæC-arm across genera; the overlap of the tRNA Pro gene with tRNA Thr on opposing strands in *Asphondylia*; the overlap of tRNA Trp with the cox1 gene in *Asphondylia*; tRNA Tyr shown here in Mayetiola is found in a different genomic positions in each genus, between nad2 and cox1 in *Rhopalomyia*, between nad3 and nad5 in *Mayetiola* and between cytb and nad1 in *Asphondylia*.

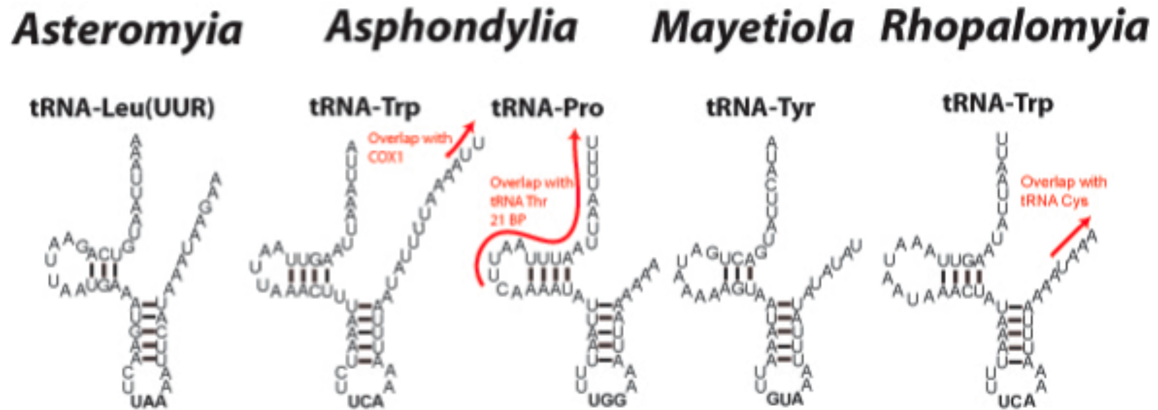


Figure 18 Phylogeny of cecidomyiid genera and outgroup based on sequence data from cytochrome oxidase subunit I gene. Numbers above branches are MP bootstrap, ML bootstrap, and Bayesian posterior probabilities. Taxa displaying incomplete inferred secondary tRNA structure are in denoted at the
tips.

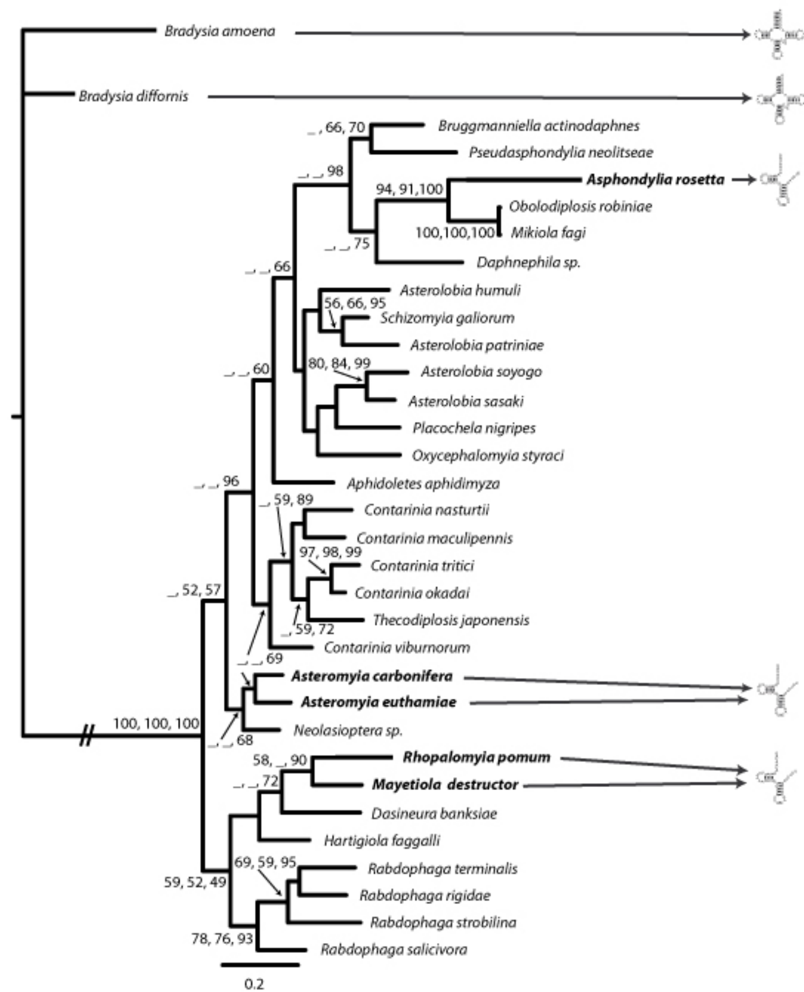


Figure 19 All tRNA genes found within the mitochondrial genome of *Mayetiola destructor.*

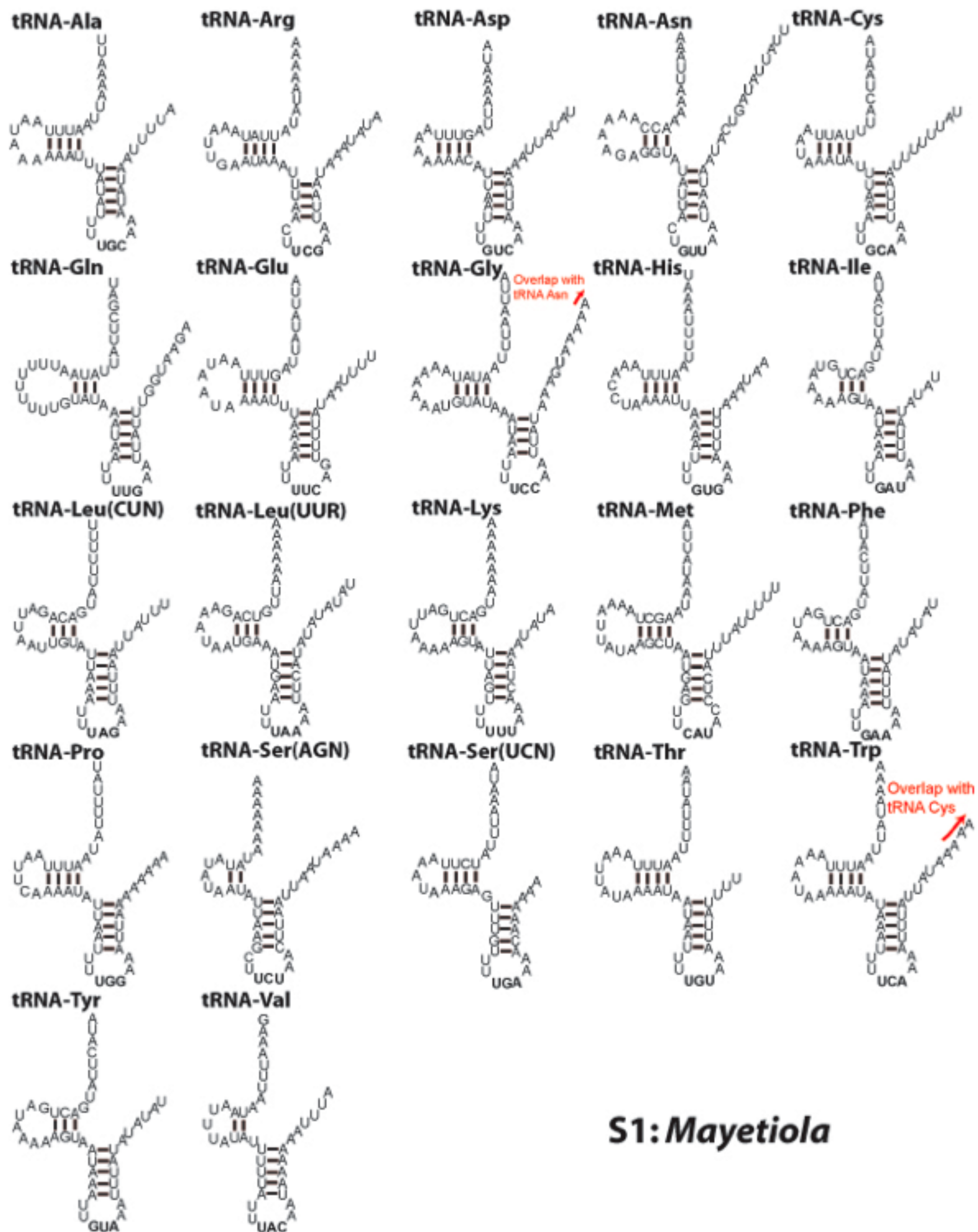


Figure 20 All tRNA genes found within the mitochondrial genome of *Rhopalomyia pomum.*

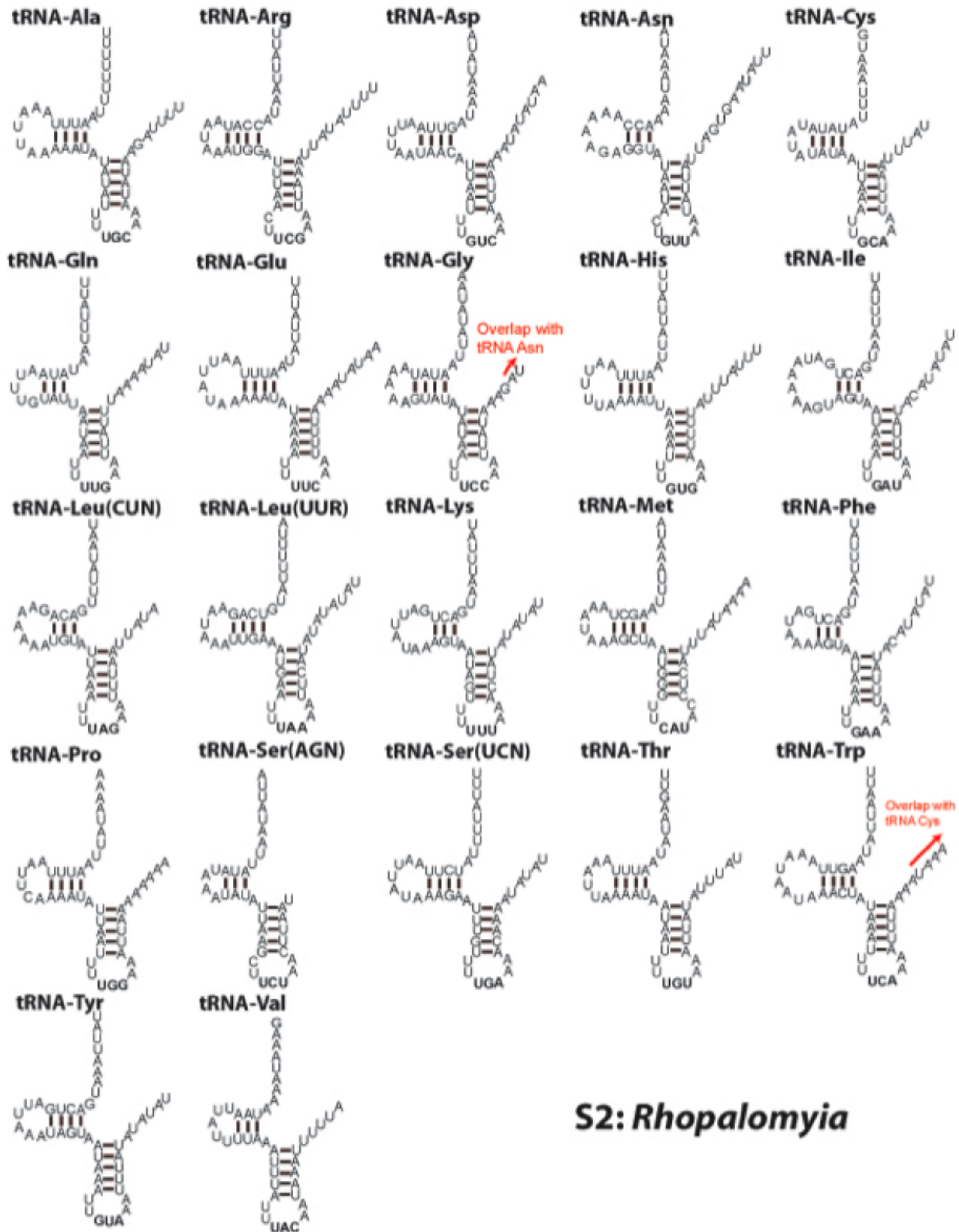


Figure 21 tRNA genes found within the mitochondrial genome of *Asphondylia rosetta* and *Asteromyia spp.*