

**ECOLOGICAL GENETICS ASPECTS OF ANTHROPOGENIC
HOST-SHIFTS IN SOAPBERRY BUGS (RHOPALIDAE)**

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Biology
University of Saskatchewan
Saskatoon

By

Prasobh Raveendran Thampy

PERMISSION TO USE

Title of thesis: Ecological genetics aspects of anthropogenic host-shifts in soapberry bugs (Rhopalidae)

Name of author: Prasobh Raveendran Thampy, Department of Biology.

Degree: Master of Science

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology

University of Saskatchewan, Saskatoon, Saskatchewan (S7N5E2)

ABSTRACT

Host range expansion, or adaptation of insects to new hosts, is a worldwide phenomenon that has been observed repeatedly and extensively; still, the genetic mechanisms behind host-shifts are not well known. In this thesis I focus on the morphological and genetic variation associated with two recent anthropogenic host shifts in two species of soapberry bug, *Leptocoris tagalicus* and *Jadera haematoloma*. First, I investigated the host-associated genetic differentiation in Australian *Leptocoris* soapberry bugs, as determined by genome-wide variation patterns. My results show that specimens feeding on two naturalized Neotropical balloon vines, (*Cardiospermum halicacabum* and *C. grandiflorum*) have longer “beaks” than those living on the native trees *Atalaya hemiglauca* and *Alectryon tomentosus*. Genetic analyses of mitochondrial haplotypes and amplified fragment length polymorphic (AFLP) markers indicate that the lineage of bugs on the annual vine *C. halicacabum*, is intermediate between two subspecies of *L. tagalicus* found on the native hosts. Moreover, where this annual vine and whitewood tree (*A. hemiglauca*) co-occur, the morphology and genomic composition of the bugs are similar to those occurring in allopatry. These results show that hybridization provided the genetic elements underlying the strongly differentiated ‘*halicacabum* bugs’. In contrast, the bugs feeding on the recently introduced perennial balloon vine (*C. grandiflorum*) showed no evidence of admixture, and are genetically indistinguishable from the nearby populations on a native host.

Second I used a candidate gene approach to investigate the molecular genetic basis of host-adaptation in Floridian populations of the red-shouldered soapberry bug,

J. haematoloma. While in Southern Florida soapberry bugs have long beaks to penetrate the large fruits of the native balloon vine (*C. corindum*), in northern and central Florida, bugs have evolved to feed on an introduced, flat-podded host, the Taiwanese Golden rain tree (*Koelreuteria elegans*). Specifically I focused on five genes because of their potential role in host preference (*orco*), “beak” length (*Dll*, *dac*, *hth*) and the adaptation to the toxic compounds of host-plants (Na^+/K^+ -ATPase). My results suggest these genes are highly conserved in this system, and that genetic variation at these loci is not associated with the different host-plants.

ACKNOWLEDGMENTS

I would like to thank the Department of Biology, University of Saskatchewan and Dr. José Andrés for giving me the opportunity to learn and carry out the research. Many thanks go to the members of my advisory committee for the support during my pursuit of the degree. I am also very thankful to all my lab mates, Pam, Halyna and Gillian. I would like to thank Scott P. Carroll for his comments, insight, and discussions regarding the research, and his help with sample collection. I greatly appreciate the department, for scholarship support and the Discovery Grant to JAA (Natural Sciences and Engineering Research Council of Canada) for funding of the research.

I would like to use this opportunity to thank Sudheej, Teenus and their family for their unconditional love and support. I am also thankful to all my friends for their constant supports. Lastly but most importantly, without the unconditional love and support from my family. I could not have made it this far, I am forever grateful to all.

TABLE OF CONTENTS

PERMISSION TO USE	i
ABSTRACT	ii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	vii
LIST OF TABLES	ix
CHAPTER 1- Introduction	1
1.1. Objectives	4
CHAPTER 2- Population genetics of a host-shift in Australian <i>Leptocoris</i>	6
2.1. Introduction	6
2.2. Materials and methods	9
2.2.1. <i>Leptocoris</i> identification, host use and morphological Measurements	9
2.2.2. Host plants: Sampling and morphological measurements	13
2.2.3. Sampling for genetic analyses	14
2.2.4. <i>Leptocoris</i> mitochondrial DNA amplification and sequencing	17
2.2.5. Fluorescent Amplified Fragment Length Polymorphism Methods	17
2.2.6. Phylogenetic reconstruction: Mitochondrial and nuclear markers	18
2.2.7. Population structure	20
2.2.8. Genome-wide admixture estimates and hybrid identification	22
2.3. Results	23
2.3.1. Host plants: Distribution and morphological measurements	23
2.3.2. <i>Leptocoris</i> morphological measurements	24
2.3.3. Phylogenetic reconstruction	25
2.3.4. Population structure	30
2.3.5. Genome-wide admixture estimates and hybrid identification	32
2.4. Discussion	33
2.5. Conclusions	42

CHAPTER 3- Genetics of host-shift and adaptation in <i>Jadera</i> bugs of Florida, using candidate gene approach	43
3.1. Introduction	43
3.2. Materials and methods	48
3.2.1. Characterization of the candidate genes	48
3.2.2. Identification of genetic variation among populations	50
3.2.3. Phylogenetic analyses	50
3.2.4. Population genetic analyses	52
3.3. Results and discussions	52
3.3.1. Sequence analyses	52
3.4. Conclusions	55
CHAPTER 4- Conclusions and future directions	58
REFERENCES	60

LIST OF FIGURES

<u>Figure</u>	<u>page</u>	
1.1	Overview of the Heteropteran mouthparts. (A) Morphology of the beak of the soapberry bug. (B) Mouthparts separated to show specific anatomical features including the slender four-segmented labium, two pairs of stylets and labrum.	2
2.1.	Ecological niche modeling estimates of suitable habitats and the geographical overlap between host plants.	16
2.2.	Ecological niche modeling of different native and introduced host plants utilized by <i>Leptocoris tagalicus</i> . A. whitewood (<i>At. hemiglauca</i>). B. Woolly Rambutan (<i>Al. tomentosus</i>). C. Annual Balloon Vine (<i>C. halicacabum</i>). D. Perennial Balloon Vine (<i>C. grandiflorum</i>).	26
2.3.	A) <i>At. hemiglauca</i> / <i>Al. tomentosus</i> overlap, B) <i>C. grandiflorum</i> / <i>Al. tomentosus</i> overlap and C) Overlap between <i>Al. hemiglauca</i> , <i>Al. tomentosus</i> and <i>C. grandiflorum</i> .	27
2.4.	Box-plots of the grand means of morphological differences between both <i>Leptocoris</i> subspecies, and different host lineages. Boxes represent the first and third quartiles, whiskers represent 5 and 95% percentiles. Population means outside this range are represented as dots.	29
2.5.	1, Bayesian tree and 2, ML tree. Colors represent different host plants, Introduced: <i>C. halicacabum</i> (purple), <i>C. grandiflorum</i> (light blue). Native: <i>At. hemiglauca</i> (red), <i>Al. tomentosus</i> .	35
2.6.	1) Map of Australia showing the locations of the populations used in the genetic analyses. Colors represent different host plants. Native: <i>At. hemiglauca</i> (red) and <i>Al. tomentosus</i> (yellow); Introduced: <i>C. halicacabum</i> (purple) and <i>C. grandiflorum</i> (light blue), while different shapes mark different populations within each host. 2) Median-joining network describing the evolutionary relationships between mtDNA haplotypes of <i>Leptocoris tagalicus</i> . Each haplotype is represented by a circle whose area reflects the overall number of copies observed and whose color-coding indicates the populations where the haplotype is found. Line length is proportional to the number of differences between haplotypes (tick marks). Small red circles represent non-sampled ancestral haplotypes that were reconstructed by the MJ algorithm.	36

<u>Figure</u>	<u>page</u>
2.7.	Phylogenetic structure of <i>L. tagalicus</i> based on ~1,200 EcoRI-MseI AFLPs. A. B. Maximum parsimony and Neighbor Joining trees. Numbers represent the bootstrap support (n=1,000). C. STRUCTURE analysis. The color of the individuals (rows) represents the proportion of their genome assigned to the K=2 and K=3 inferred clusters in the model-based admixture analyses D. Multivariate analyses (DAPC) Dots represent each individual. Groups are delineated by inertia ellipses. The analysis accounted for 60% of the genetic variation in the data set of which 26% and 19% are explained in first (x-axis) and second (y-axis) principal components. Colors represent different host plants. Native: <i>At. hemiglauca</i> (red) and <i>Al. tomentosus</i> (yellow); Introduced: <i>C. halicacabum</i> (purple) and <i>C. grandiflorum</i> (light blue), whilst different shapes mark different populations within each host. 37
2.8.	Assignment of ancestry of the annual vine bugs from two parental populations (grey dots). Black and white dots represent <i>tagalicus</i> and <i>vulgaris</i> individuals respectively. <i>vulgaris</i> individuals are further split into allopatric and sympatric populations. In sympatry, two different hosts, the non-native annual vine (<i>C. halicacabum</i>) and native whitewood tree (<i>At. hemiglauca</i>) coexist sometimes only a few meters apart. 38
2.9.	Relationships between admixture index, beak length and beak/body ratio. Black and white dots represent individuals collected on annual vine (<i>C. halicacabum</i>) and whitewood (<i>At. hemiglauca</i>) respectively. Distribution of <i>Jadera</i> populations in Florida. 39
3.1.	Schematic representation of activation of ion channel by volatile compounds (Odorants) produced by plants, Interaction between the ORs and <i>orco</i> receptors. 47
3.2.	<i>Jadera haematoloma</i> populations and sampling sites. Upper right: map of the United States of America with Florida highlighted. Left: Florida. Lower right: close-up of the Florida Keys and part of the tip of the peninsula. Black circles represent known populations of the ancestral long-beaked ecomorph feeding on the native balloon vine (<i>C. corindum</i>). Stars represent known populations of the derived short-beaked ecomorph feeding on the introduced golden rain tree (<i>K. elegans</i>). 51
3.3.	Phylogenetic trees A, Maximum Likelihood tree, B- Neighbor Joining tree. Red color represents Golden rain tree population and blue represents <i>Cardiospermum</i> populations. 56
3.4.	Multiple sequence alignment of Na ⁺ /K ⁺ -ATPase (α -subunit) of representative sequences from different insect taxa. * Represents sequence from insects feeding on cardenolides containing seeds. <i>Jadera</i> * represents the gene fragment sequenced from a RNA pool of bugs populations feeding on cardenolides containing host plant. Full-length sequences are given with their total number of amino acid residue numbers. Small hexagon symbols on the top of alignment block represents the amino acid substitutions involved in cardenolide binding. 57

LIST OF TABLES

<u>Table</u>	<u>page</u>
2.1. Quick reference to host-associated ecotypes of <i>Leptocoris tagalicus</i> in Australia.	10
2.2. Represents <i>Leptocoris</i> sampling site and their respective hostplants.	12
2.3. Beak length differences in <i>Leptocoris tagalicus</i> . Linear mixed effect model including sex and host plant as fixed effects and sampled populations as random variable. The significance of the fixed factors was estimated using likelihood ratio tests between the Akaike Information Criteria (AIC) of the full model with those of the nested models excluding the appropriate variable (df, degrees of freedom).	28
3.1. List of Primers.	49
3.2. List of <i>Jadera</i> populations from Florida and their associated hosts.	49

CHAPTER 1 – INTRODUCTION

Adaptation is the most important evolutionary process in creating biodiversity, and is governed by the environment, natural selection, genes, development, and phenotypic variations. Understanding how these factors interact to produce adaptive change can be difficult for any natural study system, making it hard to address important questions about how the organisms adapt to their environment. However, rapid responses of organisms to anthropogenic changes such as non-native species invasions, land use conversion, climate warming, landscape alteration, agricultural intensification and spread of pathogens provide us with direct opportunities to study adaptation-in-action. Indeed, phenotypic changes associated with human-disturbed environments are often greater than those precipitated by ‘natural’ agents and are often adaptive (Carroll 2007b; Ghalambor *et al.* 2007; Hendry *et al.* 2008; Palumbi 2001; Strauss *et al.* 2006).

Biological invasions can generate strong signals of selective change (Keller & Taylor 2008; Lambrinos 2004; Lavergne *et al.* 2010; Simon & Townsend 2003) and are ideal systems for studying contemporary evolution and its consequences. Beak length evolution in soapberry bugs (*Rhopalidae*) in response to a novel, introduced host plant represents an exceptional model to study the entire adaptive recursion from ecology to genes to development to biodiversity. These insects feed on seeds of the Sapindaceae or soapberry plants family. To do so, they use their beak-like structures (Fig 1.1) to reach the seeds inside the fruit. Because host plants differ in their fruit size the beak size of each species varies in length. Moreover, in those species of bugs feeding on more than one host, the beak length of the different populations has increased or decreased in the direction predicted by the fruit size of the host.

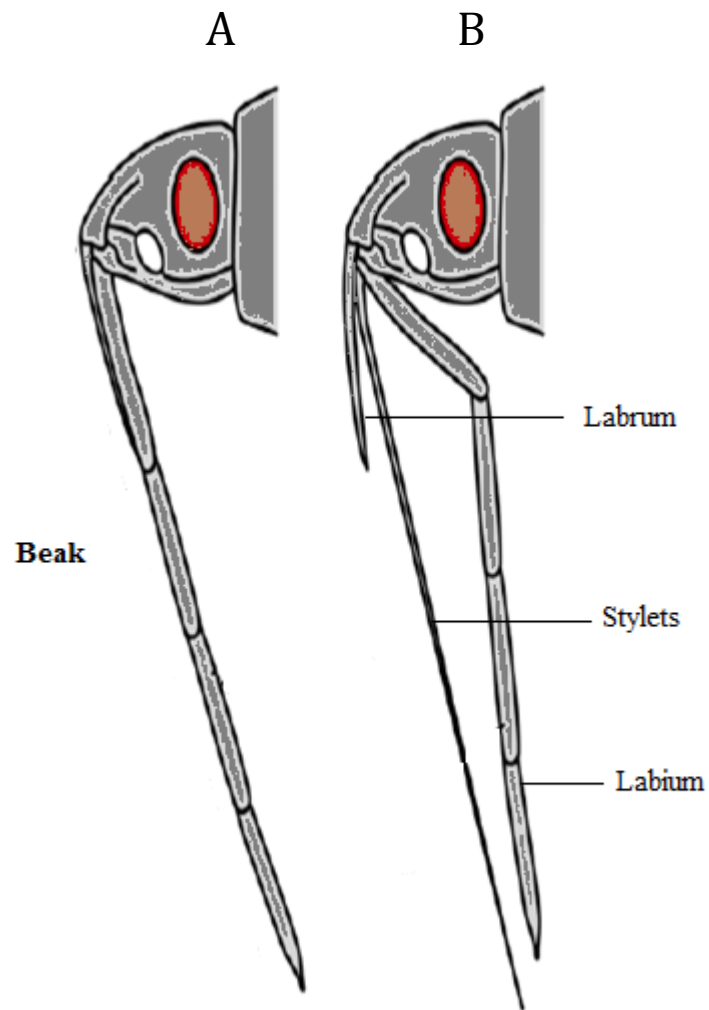


Figure 1.1. Overview of the Heteropteran mouthparts. (A) Morphology of the beak of the soapberry bug. (B) Mouthparts separated to show specific anatomical features including the slender four-segmented labium, two pairs of stylets and labrum

In Florida, the red-shouldered soapberry bug (*Jadera haematoloma*) feeds on the seeds of both a native balloon vine (*Cardiospermum corindum*) and the recently introduced (~50-60 years ago) Taiwanese golden rain tree (*Koelreuteria elegans*). While in Southern Florida soapberry bugs have long beaks to penetrate the large fruits of the native balloon vine, in northern and central Florida bugs have evolved to feed on the flat-podded introduced tree and possess much shorter beaks (Carroll & Boyd 1992; Carroll *et al.* 2003; Carroll *et al.* 1998; Carroll *et al.* 2005a). In parallel, changes have also occurred in development time, clutch size, egg mass, body size and shape, and host preference. As a result, in less than 80 years populations of this species have diverged into a native and a derived ecomorph (Carroll *et al.* 2001; Carroll *et al.* 1998; Carroll *et al.* 2005a; Dingle *et al.* 2009).

Meanwhile, a series of similar evolutionary changes have been playing out in Australia, with two species of Neotropical balloon vines as invaders rather than natives. In this case, the Australian soapberry bug (*Leptocoris tagalicus*) has colonized the vines, including the invasive perennial balloon vine (*C. grandiflorum*) in eastern forests, which have much larger fruits than native hosts. Driven by selection for foraging efficiency, in ~40 years, bugs adopting this host have evolved longer beaks than those on feeding on adjacent native *Alectryon* trees. This phenotypic change has increased the kill rate of exotic vine seeds by half (Carroll *et al.* 2005b, 2006). In northern Australia, a second, even more derived *L. tagalicus* population exists on a rare annual vine (*C. halicacabum*) that has been present for at least 200 yrs. Compared to bugs on neighboring small-fruited native *Atalaya* trees, beak length of the northern bugs has increased. As a result, bugs feeding on the annual vine have a much higher beak/body length ratio, isometric to that

of the efficient Neotropical soapberry bugs of the *Leptocoris* genus (Carroll & Loye 2006).

Investigating different genetic aspects of how this evolution has taken place, is the goal of this thesis. Although there are numerous phenotypic and genetic changes associated with host shifts in soapberry bugs, for the most part I focused on beak length because of the beak's clear adaptive value for bugs to reach the seed resources from outside the capsule before dehiscence (Carroll *et al.* 2003), and because of the beak's direct causal link with the biological control value of the derived Australian populations of *Leptocoris* (Carroll *et al.* 2005b).

1.1. Objectives

Ecological relevant traits such as beak length in phytophagous heteropterans are likely to converge across distinct lineages if they are repeatedly favored by natural selection in similar environments (e.g.,(Butlin *et al.* 2008; Hubbs 1940; Langerhans *et al.* 2007; Ley & Hardy 2014; Schluter & Nagel 1995)). Although closely related species might utilize different genetic pathways to develop the same phenotypes (Arendt & Reznick 2008), several examples show that this is not always the case and identical genetic processes often underlie similar phenotypic changes (Arendt & Reznick 2008; Hoekstra *et al.* 2006; Rosenblum *et al.* 2004). To the best of my knowledge, there is no molecular information regarding the genetic basis underlying the phenotypic changes associated with host-shifts in soapberry bugs. This thesis represents the first step in understanding the genetic basis of rapid beak-length evolution in response to anthropogenic introduction of new hosts.

Specifically, I have two objectives: **First**, to identify the genetic variation pattern behind the anthropogenic host-shifts in *Leptocoris tagalicus* associated with the colonization of two naturalized balloon vines in Australia. **Second**, to identify and characterize a set of candidate genes involved in the colonization of the new hosts in *Jadera haematoloma*. To achieve my objectives, I combined molecular and population genetics approaches with geographic information systems, phylogeography and bioinformatics tools.

CHAPTER 2- POPULATION GENETICS OF A HOST-SHIFT IN AUSTRALIAN *LEPTOCORIS* BUGS

2.1 Introduction

Plant-feeding insects, together with their hosts, comprise close to half of Earth's known species diversity (Futuyma & Agrawal 2009). A large proportion of these insects feed on just one or a few species of closely related plants (Novotny *et al.* 2004), and many species show highly specialized adaptations to exploit their particular hosts (Despres *et al.* 2007; Karban & Agrawal 2002; Toju 2009). These patterns support the notion that the great diversity of phytophagous insects stems from their ability to colonize and adapt to new plant hosts. Evolutionary biologists have long mined species level phylogenies to examine hypotheses for insect-plant diversification, and it is now widely recognized that host shifts play a key role during speciation events in plant-feeding insects (Ehrlich & Raven 1964; Janz *et al.* 2006; Mitter *et al.* 1988; Nyman *et al.* 2010). Phylogenetic analyses alone provide little insight about the genetic changes that permit successful host colonization and host adaptation (Via 2009), for which a more suitable approach is to focus on the microevolutionary aspects of new host associations. Insect populations colonizing novel hosts sometimes differentiate rapidly from locally-adapted populations that still use the original host species. This rapid change permits the study of the ecological and genetic bases of host shifts in recent and contemporary time (e.g., (Carroll *et al.* 2003; Carroll *et al.* 2001; Dingle *et al.* 2009; Feder *et al.* 2003a; Peccoud *et al.* 2009)).

¹ This chapter has been peer reviewed and published in *Molecular Ecology* (Andres, et al., 2013) DOI: 10.1111/mec.12553. I generated all the molecular data, carried out the analyses and wrote the draft of the manuscript.

Introduced, naturalized plants often attract one or more resident herbivores (Carroll & Loye 2012; Frenzel & Brandl 2003; Graves & Shapiro 2003), and because colonization events can be monitored close to the time of hosts' introduction, they have proven to be key to unraveling the genetic origins of new lineage formation (e.g. (Bush 1969; Carroll *et al.* 1997; Carroll *et al.* 2005a; Feder *et al.* 2003a; Gompert *et al.* 2006; Schwarz *et al.* 2005; Via 1991a, b)). Several genetic mechanisms for the rapid differentiation on novel hosts have been identified. Those include not only new mutations, but also large non-additive effects within existing genetic variation (Carroll 2007a; Carroll *et al.* 2003) and the introgression, through hybridization, of pre-adapted genetic elements that originated outside of currently adapting populations (Feder *et al.* 2003a). After the initial colonization of the new host, barriers to gene flow might then evolve as a result of ecologically-based divergent selection on new versus old host plants (Rundle & Nosil 2005; Schluter 2009).

In this chapter, we test if hybridization between ecologically divergent lineages is associated with phenotypic novelty in two different anthropogenic host shifts. Soapberry bugs are Rhopalid true bugs specialized as seed predators of plants of the Sapindaceae family throughout much of the world (Carroll & Loye 2012). The Australian soapberry bug *Leptocoris tagalicus* Burmeister has shifted onto two naturalized Neotropical balloon vine (*Cardiospermum*) species in Australia that are regarded as environmental weeds. The inflated fruits of these balloon vines, which have evolved in concert with the beak length of their New World *Jadera* seed predators (Carroll & Loye 1987), are much larger than the fruits of any native Australian hosts. Gross (1960) described two morphologically distinct subspecies of *L. tagalicus* in Australia, and these are associated with different native trees that have

largely separate geographic distributions (Carroll *et al.* 2005a; Carroll *et al.* 2005b; Gross 1960). In recent decades, bugs of the eastern subspecies that have shifted to *C. grandiflorum* have evolved beaks averaging about 10% longer than on the native tree as a result of selection for foraging efficiency on seeds in the novel inflated fruits (Carroll *et al.* 2005a). The second subspecies is smaller, with a shorter beak and feeds on small-fruited native trees in desert and monsoonal Australian habitat. In addition, a very long-beaked small morph has recently been discovered in patches of balloon vine (*C. halicacabum* v. *halicacacabum*; these plants were likely introduced more than 200 years ago (Bean 2007). Host-associated differences in beak length are heritable and are associated with the frequency with which seeds are reached by foraging bugs (Carroll *et al.* 2005a).

Based on genetic findings we present here, we propose that the bugs on *C. grandiflorum* are derived directly from populations on native eastern hosts, but that bugs on *C. halicacabum* (hereinafter ‘Halicacabum bugs’) result from hybridization between the two subspecies. Their hybridization is associated with a novel phenotype with long beaks and high beak length/body length ratios that circumvent the inflated seed defense of the annual balloon vine that has coevolved with New World soapberry bugs but is unprecedented in Australia.

2.2 Materials and Methods

2.2.1 *Leptocoris* identification, host use and morphological measurements

There are two lineages of *L. tagalicus* that were described by Gross (1960) as the subspecies *vulgaris* and *tagalicus* (Table 2.1). *Leptocoris tagalicus tagalicus* (hereinafter *tagalicus*) is a relatively large bug that populates eastern forests, where it principally feeds on the native Woolly Rambutan tree (*Alectryon tomentosus*) and the Neotropical *perennial* balloon vine (*C. grandiflorum*), an early 20th Century horticultural introduction that became widely naturalized in the 1960s (Carroll *et al.* 2005a). In contrast, *Leptocoris tagalicus vulgaris* (hereinafter *vulgaris*) is a smaller bug that mostly feeds on the flat seeds of the native whitewood tree (*Atalaya hemiglauca*) in inland deserts and savannahs as well as seasonally inundated northern floodplains. In these northern regions, we also found a third form of *L. tagalicus* feeding on seeds in the inflated fruits annual balloon vine (*C. halicacabum*), a Neotropical native that apparently arrived between 1600 and 1800 (perhaps from 16th Century Portuguese introductions to Ambon, Indonesia (Bean 2007), although an earlier introduction date is also possible. While bugs on this host are morphologically distinct (see Table 2.1) we initially postulated that this form was derived from *vulgaris* bugs on *Atalaya* due to their geographic proximity and smaller body size.

In total, we collected bug morphological data from 57 populations (Table 2.2). Individuals were identified and assigned all collected specimens following the descriptions and illustrations of Gross (1960). We sexed each individual and measured the body length, pronotum width, and beak (labial) length using digital calipers as described in Carroll *et al.* (2005a).

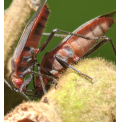



Subspecies	Geographic range	Host species	Host provenance	Relative fruit size	Relative beak length	Relative body length	Insects on host fruits
<i>tagalicus</i>	Eastern forests	<i>Alectryon tomentosus</i> and congeners	Native	Medium	Medium	Large	
<i>tagalicus</i>	Eastern forests	<i>Cardiospermum grandiflorum</i>	Introduced early 1900s, widespread and common since 1960s	Very large	Medium-large	Large	
<i>vulgaris</i>	Mainly interior of N half of continent	<i>Atalaya hemiglauca</i> and congeners	Native	Small	Small	Small	
halicacabum bug	Northern areas of NT	<i>Cardiospermum halicacabum</i>	Introduced > 200 y	Large	Large	Small-medium	

Table 2.1. Quick reference to host-associated ecotypes of *Leptocoris tagalicus* in Australia.

Host Plant	Location	Population Code
<i>Alectryon tomentosus</i> *	<i>Al T-Bahr 1</i>	<i>ABR</i>
<i>Alectryon tomentosus</i>	<i>Al T-Bahr 2</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Bahr 3</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Bnlgh</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Coot</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Lgnhlm</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Nebo</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Olmeau</i>	
<i>Alectryon tomentosus</i>	<i>Al T-Raft</i>	
<i>Alectryon tomentosus</i> *	<i>Al T-Sherwood 1.2</i>	<i>SHP</i>
<i>Alectryon tomentosus</i>	<i>Al T-Sherwood 1.1</i>	
<i>Alectryon tomentosus</i> *	<i>Al T-Thompson rd</i>	<i>THR</i>
<i>Atalaya hemiglauca</i> *	<i>AH-Alice 1</i>	<i>AAS</i>
<i>Atalaya hemiglauca</i> *	<i>AH-Alice 2</i>	<i>ASP</i>
<i>Atalaya hemiglauca</i>	<i>AH-Alice 4</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Alice 5</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Alice 6</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Clermont</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Goond</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Injune</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Lynd</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Marlborough</i>	
<i>Atalaya hemiglauca</i> *	<i>AH-Mitchell</i>	<i>AMA</i>
<i>Atalaya hemiglauca</i> *	<i>AH-Mt. Elliot Station</i>	<i>AES</i>
<i>Atalaya hemiglauca</i> *	<i>AH-NewH</i>	<i>ANH</i>
<i>Atalaya hemiglauca</i> *	<i>AH-Roma 1</i>	<i>RMA</i>
<i>Atalaya hemiglauca</i>	<i>AH-Roma 2</i>	
<i>Atalaya hemiglauca</i>	<i>AH-Roma 3</i>	
<i>Atalaya hemiglauca</i> *	<i>AH-Victoria-1.2</i>	<i>VRC-a</i>
<i>Cardiospermum grandiflorum</i>	<i>CG-Atherton</i>	
<i>Cardiospermum grandiflorum</i>	<i>CG-Boonora</i>	
<i>Cardiospermum grandiflorum</i>	<i>CG-Bremer</i>	
<i>Cardiospermum grandiflorum</i>	<i>CG-Bundaberg</i>	
<i>Cardiospermum grandiflorum</i>	<i>CG-Chinchilla</i>	
<i>Cardiospermum grandiflorum</i>	<i>CG-Cow</i>	

<i>Cardiospermum grandiflorum</i>	CG-Gayn	
<i>Cardiospermum grandiflorum</i>	CG-Goond	
<i>Cardiospermum grandiflorum</i>	CG-KhoBot	
<i>Cardiospermum grandiflorum</i>	CG-KhoBridge	
<i>Cardiospermum grandiflorum</i>	CG-Load	
<i>Cardiospermum grandiflorum</i>	CG-Long Pocket	
<i>Cardiospermum grandiflorum</i>	CG-Lowood	
<i>Cardiospermum grandiflorum</i>	CG-Mt Cros2	
<i>Cardiospermum grandiflorum</i>	CG-Mt Crosby 1	
<i>Cardiospermum grandiflorum</i>	CG-Nebo	
<i>Cardiospermum grandiflorum</i>	CG-Nerang	
<i>Cardiospermum grandiflorum</i>	CG-Shep	
<i>Cardiospermum grandiflorum</i>	CG-Sherwood	
<i>Cardiospermum grandiflorum</i> *	CG-Doug Larsen park	DLP
<i>Cardiospermum grandiflorum</i> *	CG-St Lucia Golf	ESL
<i>Cardiospermum grandiflorum</i> *	CG-Indooroopilly Golf club	IGC
<i>Cardiospermum grandiflorum</i> *	CG-Dundas	DUN
<i>Cardiospermum halicacabum</i> *	CH-Allig 1.1	EAR
<i>Cardiospermum halicacabum</i>	CH-Allig 1.2	
<i>Cardiospermum halicacabum</i> *	CH-Daly	DRC
<i>Cardiospermum halicacabum</i> *	CH-Vict 1.1	VRC-b
<i>Cardiospermum halicacabum</i>	CH-Vict 1.3	

Table 2.2. Represents *Leptocoris* sampling site and their respective hostplants.

* denotes the samples used for genetic analyses.

For each subspecies we analyzed differences in the measured traits using linear mixed effect models, including sex and host plant as fixed effects and populations as random variable. Models were fitted in R version 2.15.2, implementing the function *lmer* (package *lme4*) with normal errors and identity link function. As a global test of the significance of fixed effects we used likelihood ratio tests to compare the overall fit of the full model with that of the appropriate nested models. Approximate P-values for each level were obtained by calculating their 95% CI using 10,000 Monte Carlo simulations with *mcmcsmpl*.

2.2.2 Host plants: Sampling and morphological measurements

For each host we measured the fruit size as the distance between the outside of an intact fruit and the center of the nearest seed. Depending on species, fruit size data were taken from 2-6 populations, with 2 to 30 total measurements taken from 1-5 individuals per population. We made all measurements with Mitutoyo digital calipers (Mitutoyo CD 6B, America Corporation, IL, USA).

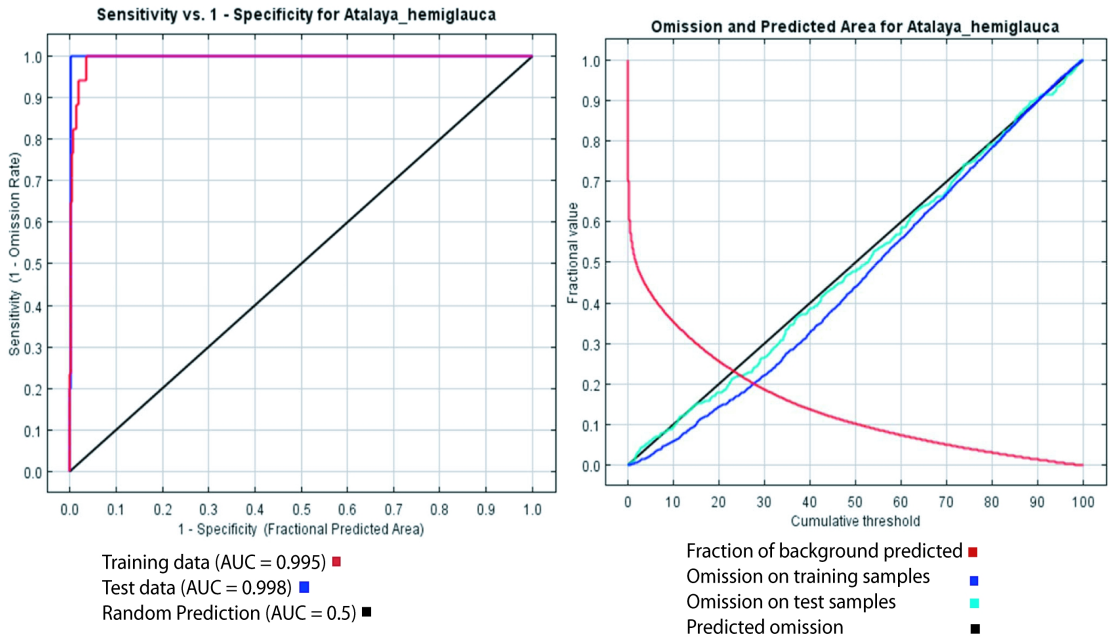
Based on collection records of the native and introduced host species, we used ecological niche modeling (ENM) to estimate the geographic distribution of suitable habitats for native and introduced hosts and infer the geographically overlap between host plants. To do so we used Maxent version 3.2 (Phillips *et al.* 2006). This machine learning approach generates the expected distribution of a species using data on the environmental conditions where it is known to occur and produces a map of the species' potential geographic distribution with a statistical likelihood of the species occurrence at each location. To construct these models, we used 19 temperature and precipitation variables from the WorldClim data set with 30-second spatial resolution

(Hijmans *et al.* 2005) and a set of georeferenced occurrence locations for native hosts *Alectryon tomentosus* (n = 251), and *Atalaya hemiglauca* (n = 2,793) as well as the introduced host vines *C. grandiflorum* (n= 267) and *C. halicacabum* (n= 324) obtained from the Australian Virtual Herbarium (avh.ala.org.au/). Locality records occurring within the same map pixel were removed to avoid pseudoreplication (details of the models are provided on Figure 2.1). To map the locations where suitable habitats for the host plant species come into geographic contact, we first classified as climatically unsuitable any grid cell falling in the lower 5th percentile (Chatfield *et al.* 2010). Then, we defined the area of overlap as the combined grid where there is least a 10% probability that the s co-occur. All calculations were done using the ‘*grid overlay*’ function of DIVA GIS 5.2 (Hijmans *et al.* 2005).

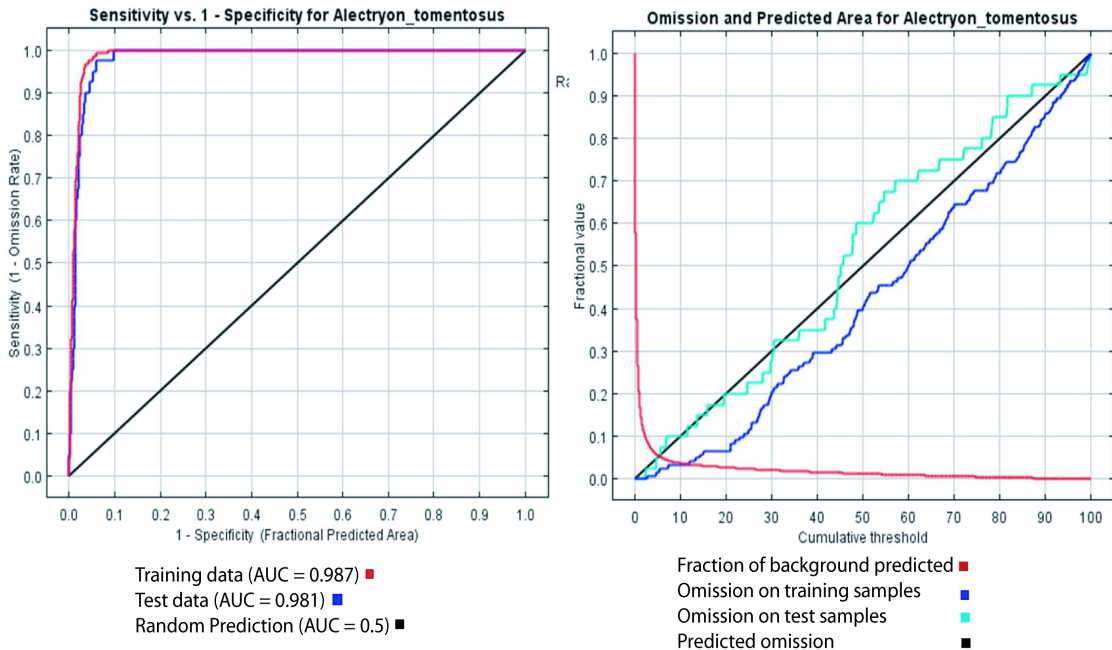
2.2.3 Sampling for genetic analyses

A subsample of 89 specimens of *tagalicus* and *vulgaris* from 17 different locations was selected for genetic analyses. These locations extend from the east coast westward to the interior and north to the Northern Territories of Australia (Table 2.1). The east coast includes bugs on the native host *Al. tomentosus*, and bugs on the introduced perennial vine *C. grandiflorum*. All these bugs are classified as *tagalicus* (Gross 1960). The Northern Territory and Central populations are classified as *vulgaris* and include the bugs on introduced *C. halicacabum*, and bugs on the native host *At. hemiglauca*. Collected individuals were preserved in ethanol at -80 °C. Genomic DNA was extracted from the thorax of the stored samples using the Epicentre MasterPure ® Kit following manufacturer's instructions.

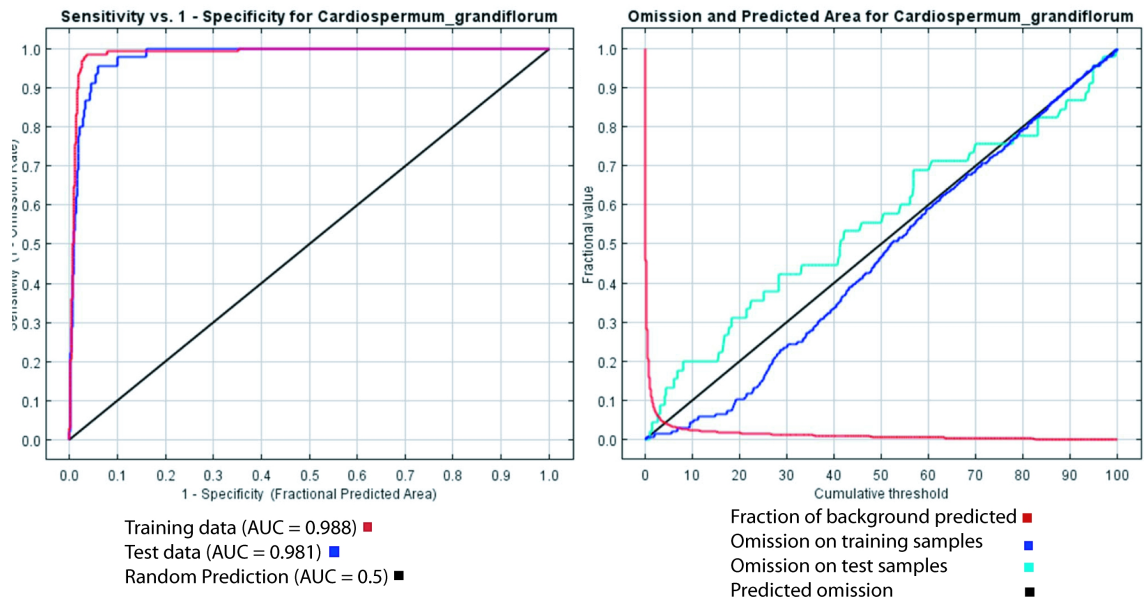
A



B



C



D

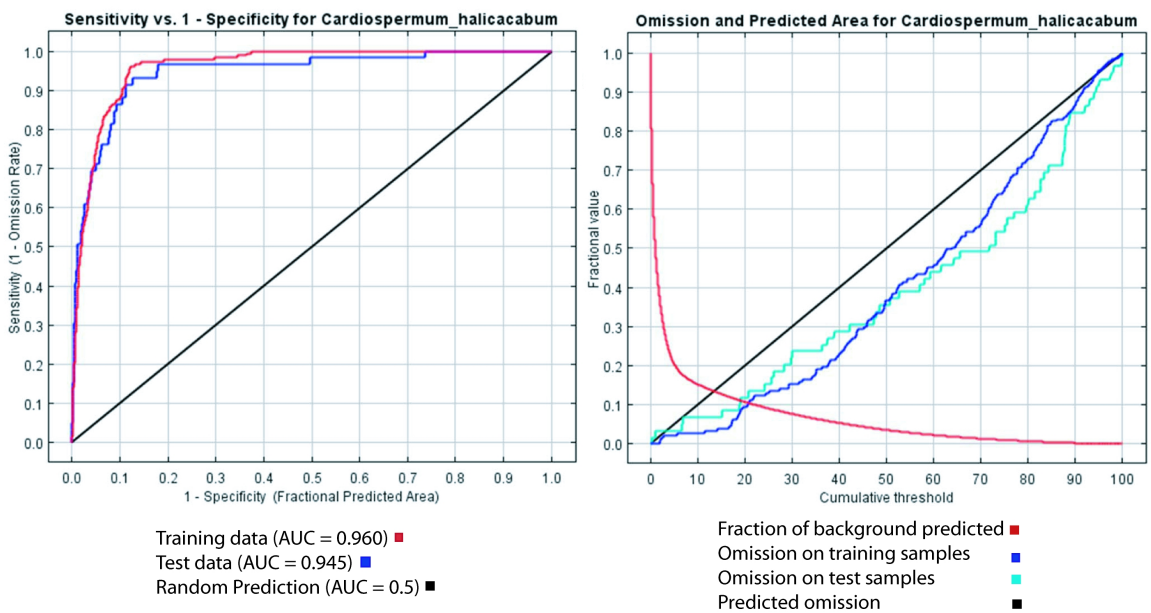


Figure 2.1. Ecological niche modelling estimates of suitable habitats and the geographical overlap between host plants.

2.2.4 *Leptocoris* mitochondrial DNA amplification and sequencing

We amplified two mitochondrial genes (*Cox-III* and *NADH-1*) using the degenerate primers (*Cox-III*: C3-J479 GTTGATTATAGACCWTGRCC and C3-N5460 TCAACAAAATGTCARTAYCA, *NADH-1*: N1-J11876 CGAGGTAAAGTMCCWCGAACCYCA and N1-N12595 GTWGCTTTTTTAACTTTATTRGARCG, (Simon *et al.* 2006) and QIAGEN Top-Taq DNA. Purified (*Exo/SAP*) PCR products were sequenced on 3130-XL Genetic Analyzer using ABI 3.1 Big Dye Terminator chemistry.

2.2.5 Fluorescent Amplified Fragment Length Polymorphism Methods

We developed fluorescent AFLP markers following (Berres 2003; Vos *et al.* 1995) with the modifications implemented in Berres (2003). Genomic DNA (100 ng) was digested with 2 U *EcoRI* and 4 U *MseI* (New England Biolabs, Ipswich, MA) for 30' at 37°C in x 1 NEB buffer 4. Then, to ligate the resulting fragments to the adapters, we added 10 µl of restriction products a 30 µl reaction mixture containing 0.5 µM *EcoRI* adapter, 5 µM *MseI* adapter, and 60 U T4 DNA ligase (New England Biolabs). After incubation (30°C for 90 min), we diluted the samples 10 times with ddH₂O, and used 2.5 µl of each sample as a template to conduct the pre-selective PCRs (x 1 PCR buffer, 0.5µM each of *EcoRI*-C combined with *MseI*-C, *MseI*-G, *MseI*-TC, *MseI*-AG primers, 0.2 mM dNTPs, and 0.5 U Top-*Taq*® Qiagen DNA polymerase. 20 cycles of 30 s at 94°C, 1 min at 56°C and 1 min at 72°C). To generate the AFLPs we diluted (1:20) these pre-amplified products and used them as the template for selective PCR amplifications (x 1 PCR buffer, 0.5µM of each primer, 0.2 mM dNTPs, and 0.5 U Top-*Taq*® Qiagen DNA polymerase) using a touchdown protocol (95°C for 3 min, 13 cycles of 30 s at 94°C, 30 s at 65°C and 1 min at 72°C, -

0.7°C/Cycle; and 12 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C). After screening, we selected seven combinations of *MseI-EcoRI*^{6-FAM} primers that generated clear, repeatable and evenly distributed bands (CGA-CAT, CCT-CAT, GTTC- CAT, TCTG-CAT, TCGC- CAT, AGAC- CAT). To prepare DNA fragments for separation by capillary electrophoresis, a sample loading solution was prepared by mixing 0.5 µL of 600-LIZ size standard* (Applied Biosystems) with 9 µL of Hi-Di Formamide, and 1 µL of 1:30 dilution of selective PCR amplification product. Samples were analyzed in ABI 3130xl genetic analyzer (Applied Biosystems). The presence or absence of fragments was scored with the software GeneMapper v.3.7 (Applied Biosystems). To assess the reproducibility and reliability of our AFLP fragments, we replicated ~ 10% of the samples. To reduce fragment size homoplasy we only scored fragments within the 150–500 bp size range (Caballero & Quesada 2010; Caballero *et al.* 2008; Paris *et al.* 2010).

2.2.6 Phylogenetic reconstruction: Mitochondrial and nuclear markers

For the mitochondrial dataset we aligned all sequences using MAFFT (Katoh & Standley 2013) and independently estimated the best model of evolution for the *Cox-III* (501 bp) and *NADH-1* (459 bp) fragments using the BIC criterion as implemented in JModeltest (Darriba *et al.* 2012). These tests revealed a HKY + I substitution model as the best fit for both fragments. Then, we used partition homogeneity test in PAUP*(4.0b10, (Swofford 2003)) to confirm that combining the fragments was appropriate. Finally, we used the resulting concatenated dataset for phylogeny reconstruction applying maximum likelihood (ML) and Bayesian methods as follows. We obtained ML trees using a heuristic search (stepwise addition; addition sequence= random with ten replicates; TBR branch swapping; MulTrees on). Then, to

evaluate the relative support of each node we generated 100 bootstrap replicates using the same parameters. For the Bayesian analysis we used MrBayes 3.2 (Ronquist & Huelsenbeck 2003). Ten Metropolis-coupled Markov chain Monte Carlo (MCMC) analyses were run twice for 4×10^6 generations and sampled every 100 generations (mcmcpc ngen = 4000000, nchains = 10, temp = 0.10, samplefreq = 100, burnin = 25%). We considered runs to have converged on stationarity when there were no trends in generation versus logL plots, potential scale reduction factors were near 1.0 for all parameters, and the average standard deviation of split frequencies was below 0.01. The Bayesian posterior probability of each node was determined as the number of post-burn-in sampled trees that contained each observed bipartition. We used two closely related species (*L. mitellatus* and *L. species 1*) as outgroups. We reconstructed the haplotype network (without outgroups) using a median-joining approach as implemented in NETWORK 4.6 (Fluxus Technologies Inc.).

For the AFLP dataset we first used a parsimony framework to estimate the signal-to-noise ratio of our data set ($n= 1,284$ loci). Then, we estimated phylogenetic relationships using three different methods 1) maximum parsimony (MP); 2) distance based (NJ), and 3) minimum evolution (ME). To assess for the adequacy of the phylogenetic signal we generated two different 1,000 randomized datasets, calculating the tree-length distribution *skewness* (e.g. (Hillis & Huelsenbeck 1992)) and the permutation tail probability (PTP, (Faith & Cranston 1991) as implemented in PAUP 4.0b8 (Swofford 1999). For the MP method we used a Wagner parsimony criterion because it assumes equal loss–gain probabilities of restriction sites, and fits AFLP data better than other criteria assuming unequal probabilities (Koopman 2005). For both distance-based trees (NJ and ME), we used a Nei-Li distance matrix (Nei & Li

1979). This method counts only shared presences and is likely more homologous than distances based on shared absences and presences (Kosman & Leonard 2005). We conducted MP and ME analyses using heuristic searches (10 replicates), 100,000 random additions, and tree bisection and reconnection (TBR) swapping.

2.2.7 Population structure

We examined the population structure of *L. tagalicus* using assignment test and principal components analysis (PCA) based approaches. First we used the Bayesian clustering algorithms implemented in STRUCTURE to assign individuals to genetic clusters by minimizing deviations from Hardy-Weinberg and linkage equilibrium. To do so, we implemented a model assuming admixture and correlated allele frequencies using 10 independent runs with of 1,000,000 Markov chain Monte Carlo (MCMC) iterations (burn-in =100,000) for a series of clusters (K) ranging from 1 (panmixia) to 17 (maximum number of localities sampled). Then, we used the statistic ΔK to select the value of K with the uppermost hierarchical level of population structure in our data (Evanno *et al.* 2005). Because the presence of broad geographic/taxonomic superstructures (e.g. *tagalicus* vs. *vulgaris*) might hide other structures at smaller spatial scales (e.g. host plant, population, etc.) we ran STRUCTURE analyses for the whole data set as well as for the populations within each subspecies separately.

Multivariate methods do not rely on explicit population genetics models (do not rely on Hardy–Weinberg equilibrium, nor do they suppose the absence of linkage disequilibrium), and may be preferable when many loci are available and the structure is subtle (Jombart 2008; Jombart *et al.* 2010; Jombart *et al.* 2008; Reeves &

Richards 2009). Thus, in addition, we inferred population subdivision using a discriminant analysis of principal components (DAPC) as implemented in the R package Adegenet v 1.3.2 (Jombart 2008; Jombart *et al.* 2010; Jombart *et al.* 2008). DAPC is a multivariate analysis that integrates principal component analysis (PCA) with discriminant analysis to summarize genetic differentiation between groups. We predicted the optimal number of clusters (populations) using the *k*-means clustering algorithm ‘find.clusters’, retaining all principal components. We calculated the Bayesian information criterion (BIC) for $K = 1-17$, where K = number of sampled populations. The optimal number of populations was identified as the one for which BIC showed the lowest value and after which BIC increased or decreased by the least amount. We then used DAPC to assign individuals into populations, retaining the number of principal components encompassing 60% of the cumulative variance.

To estimate the degree of genetic differentiation among different groups we first estimated allelic frequencies at AFLP-loci using a Bayesian approach (Zhivotovsky 1999), and then used these frequencies to determine (Lynch & Milligan 1994) F_{ST} . We obtained confidence intervals for this estimator using 5,000 random permutations. We also estimated the percentage of polymorphic loci and genetic diversity (H_j) following Lynch and Milligan (1994) as implemented on AFLPSURV 1.0 (Vekemans *et al.* 2002). We estimated linkage disequilibrium between AFLPs using ARLEQUIN 2.0 (Excoffier & Lischer 2010).

2.2.8 Genome-wide admixture estimates and hybrid identification

Because the analyses of population structure suggested the hybrid nature of the *halicacabum* populations we quantified the admixture of these bugs using two different methods. First, using the R-package INTROGRESS (Gompert & Buerkle 2009; Gompert & Buerkle 2010) we estimated the parental allele frequencies clustering in ‘pure’ populations (*vulgaris*: RMA, MIA, AST, NHS, ASL, MES; *tagalicus*: THR, IGC, DUN, SHP, DLP, ESL, BSR) using the function ‘*prepare.data*’. Then, we quantified the ancestry of each bug from the *C. halicacabum* populations by estimating the hybrid index (function ‘*est.h*’), which is an average of the genome-wide admixture for a given individual. Because of the dominant nature of the AFLP markers it is not possible to estimate interspecific heterozygosity (the proportion of an individual’s genome with alleles inherited from both parental populations). Thus, we did not categorize the admixed individuals into specific genotypic classes (*i.e.* F₁, F₂, BC) or advance-generation hybrids. Instead, we tested if the population of bugs feeding on the annual vine (*C. halicacabum*) shows a high incidence of F₁ hybrids as expected if these bugs comprise a hybrid zone maintained by continuous, substantial immigration from the *tagalicus* and *vulgaris* lineages and complete hybrid sterility. To do so, for each of these putative parental species we first calculated the expected genotypic frequencies at each AFLP locus assuming Hardy-Weinberg equilibrium, and used these frequencies to estimate the expected distribution across loci of dominant genotypes if the hybrid population was only entirely comprised of F₁ individuals. Then, we compared this distribution with the one observed in the *C. halicacabum* bugs using a permutation test (n= 1,000). Because the *a priori* assignment of pure parental populations may bias ancestry calculations we used STRUCTURE to assign individuals to either parental or hybrid classes using the

posterior probability of assignment (q) to $k=2$ clusters, were we considered that $q \geq 0.90$ indicates ‘pure *vulgaris*’, $q \leq 0.10$ indicates ‘pure *tagalicus*’, and $q [0.1-0.9]$ indicates hybrid individuals. We used this arbitrary probability threshold of 0.90 because it is likely to minimize the number of misidentified ‘pure-bred’ individuals while maximizing the efficiency of assigning hybrids (Burgarella *et al.* 2009; Vaha & Primmer 2006; Winkler *et al.* 2011). Using other two other thresholds ($q \geq 0.95$ and $q \geq 0.85$) yield very similar individual assignment results (data not shown). The statistical significance of differences in q -values between different lineages/populations was carried out using permutation tests.

2.3 Results

2.3.1 Host plants: Distribution and morphological measurements

The estimated geographic distributions for *At. hemiglauca*, *Al. tomentosus*, *C. halicacabum*, and *C. grandiflorum* show varying degrees of extent and overlap (Fig. 2.2). The area under the receiving characteristic (AUC) shows that the ecological niche models strongly discriminate between randomly selected locations across the study region and the training locations (AUC $_{AT. HEMIGLAUCA} = 0.995$, AUC $_{AL. TOMENTOSUS} = 0.987$, AUC $_{C. GRANDIFLORUM} = 0.988$, AUC $_{C. HALICACABUM} = 0.950$) and the test localities (AUC $_{AT. HEMIGLAUCA} = 0.998$, AUC $_{AL. TOMENTOSUS} = 0.981$, AUC $_{C. GRANDIFLORUM} = 0.981$, AUC $_{C. HALICACABUM} = 0.945$). The geographic distributions of climatically suitable habitats for the native hosts *At. hemiglauca* and *Al. tomentosus* show overlap and can be considered parapatric (Fig. 2.2 A, B, 2.3 A). The two introduced balloon vines show significant overlap with the native hosts (Fig. 2.2 A-D). However, while the suitable habitat for the perennial vine species (*C. grandiflorum*) shows almost complete overlap with the Woolly Rambutan (*Al. tomentosus*), the annual vine (*C.*

halicacabum) occupies a wider geographic region including a relatively small region of overlap with the native hosts, whitewood and Woolly Rambutan (Fig. 2.3 B, C). The size of fruits of the different host plants are very different (*glm*: $P < 0.0001$). The fruit radius (distance from the fruit exterior to the seed center) of the native hosts (*At. hemiglauca*: 2.49 ± 0.05 mm; *Al. tomentosus*: 4.25 ± 0.04 mm) is much smaller than that of the introduced vines (*C. halicacabum*: 7.64 ± 0.12 mm; *C. grandiflorum*: 12.40 ± 0.25 mm. All HSD tests $P < 0.0001$).

2.3.2 *Leptocoris* morphological measurements

The two subspecies of *Leptocoris* can be distinguished unambiguously on native hosts, with *tagalicus* being darker and larger (12.47 ± 0.03 mm) than *vulgaris* (10.9 ± 0.05 mm). There is marked sexual dimorphism, with females normally larger than males (body length *vulgaris*: 11.41 ± 0.07 vs. 10.40 ± 0.06 ; *tagalicus*: 13.24 ± 0.03 vs. 11.81 ± 0.02). There is a strong positive correlation between beak and body length. Larger individuals therefore have longer beaks (Spearman- $\rho = 0.81$, $P < 0.001$), but beak/body ratios are clearly similar to those found in smaller individuals.

To test whether hosts-shifts have resulted in morphological divergence of the bugs feeding on the introduced hosts we compared the morphology of bugs between host plants. In eastern Australia, *tagalicus* specimens collected from the introduced *C. grandiflorum* have on average marginally longer beaks than those found on adjacent native *Al. tomentosus* trees. This results in slightly higher beak/body ratios on the introduced host (Fig. 2.4, Table 2.3) and is consistent with the results reported by Carroll et al. (2005b) based on both contemporary field specimens and historical museum specimens. Similarly, but more profoundly, in central and northern Australia,

vulgaris individuals feeding on the introduced *C. halicacabum* have much longer beaks than the bugs on small-fruited native *Atalaya* trees, resulting in a much higher beak/body length ratios (Fig. 2.4, Table 2.3).

In Victoria River, NT, bugs on native *At. hemiglauca* co-occur with those on introduced *C. halicacabum* host only few meters away. We further characterized these sympatric bugs by comparing their morphology with that of the allopatric populations of *tagalicus* and *vulgaris*. Beak lengths of the individuals collected on the two sympatric hosts showed striking differences (overall $F_{3,65} = 59.0$, $P < 0.001$, Fig. 2.4). Specimens on native *At. hemiglauca* showed short beaks similar to those found in central/eastern Australia where the introduced *C. halicacabum* is absent (Tukey- $HSD = -0.015$, $P = 0.99$). Individuals on this latter host showed longer beaks of than those characterizing either *vulgaris* or *tagalicus* (Tukey- $HSD_{VULGARIS-TAGALICUS} = 1.73$, $P < 0.001$; Tukey- $HSD_{VULGARIS-HEMIGLAUCA} = 2.19$, $P < 0.001$). Similarly, the overall size of the individuals collected on introduced *C. halicacabum* (pronotum: 2.75 ± 0.050 mm, body length = 11.2 ± 0.10) was intermediate between *tagalicus* (3.13 ± 0.039 , 12.78 ± 0.18) and *vulgaris* (2.65 ± 0.058 , 9.75 ± 0.23). Thus, *Halicacabum* bugs showed much higher beak/body ratios than those observed in bugs on native *At. hemiglauca* ($F_{3,50} = 31.08$, $P < 0.001$, Fig. 2.4).

2.3.3 Phylogenetic reconstruction

In the concatenated dataset we identified 49 polymorphic sites leading to the definition of 32 haplotypes most of them being singletons. Shared haplotypes represented only 22% (n=7) of the total number of individuals. Shared haplotypes were on introduced *C. halicacabum*. Overall genetic diversity (H) was high (0.988).

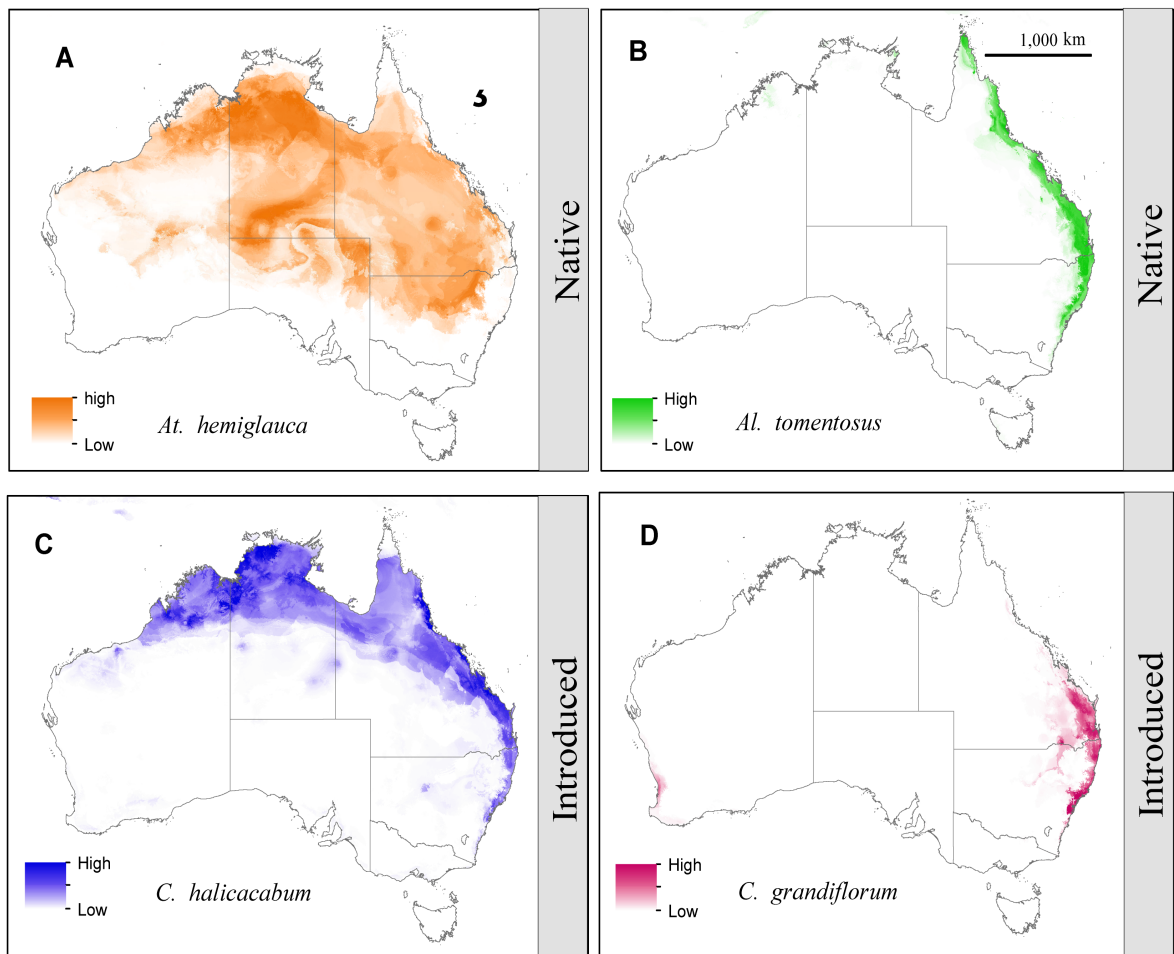


Figure 2.2 Ecological niche modeling of different native and introduced host plants utilized by *Leptocoris tagalicus*. A. whitewood (*At. hemiglauca*). B. Woolly Rambutan (*Al. tomentosus*). C. Annual Balloon Vine (*C. halicacabum*). D. Perennial Balloon Vine (*C. grandiflorum*).

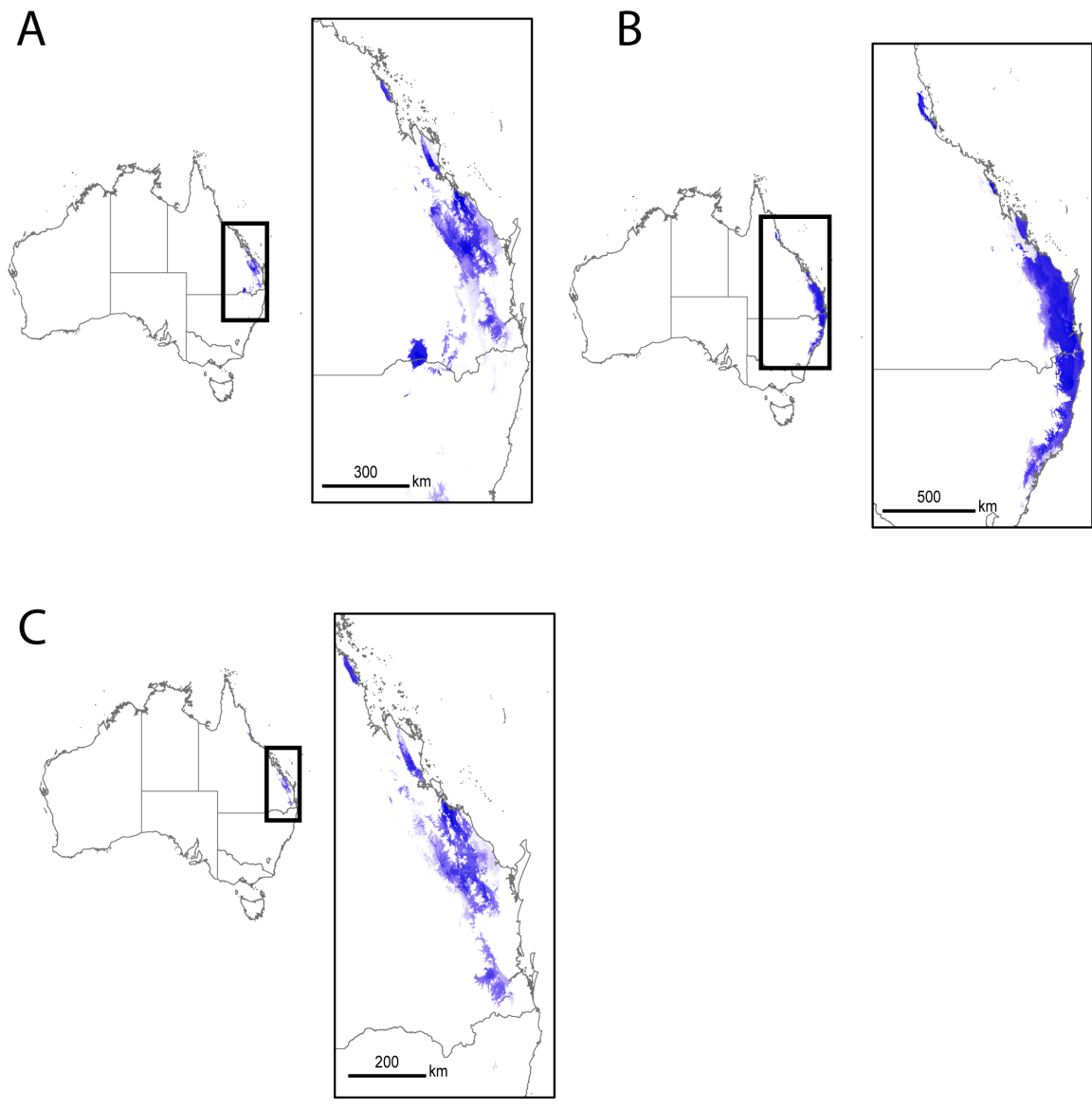


Figure 2.3. A) *At. hemiglauca* / *Al. tomentosus* overlap, B) *C. grandiflorum* / *Al. tomentosus* overlap and C) Overlap between *Al. hemiglauca*, *At. tomentosus* and *C. grandiflorum*.

Trait	Lineage	Nested Models	AIC	df	P
Beak length	<i>L. t. tagalicus</i>	Full model	15945	8	-
		+ Sex	15990	7	< 0.001
		+ Host plant	17051	4	< 0.001
	<i>L. t. vulgaris</i>	Full model	5296	5	
		+ Sex	5328	4	< 0.001
		+ Host plant	5442	4	< 0.001
Beak body ratio	<i>L. t. tagalicus</i>	Full model	12431	8	-
		+ Sex	12428	7	<0.05
		+ Host plant	12434	4	ns
	<i>L. t. vulgaris</i>	Full model	4815	5	-
		+ Sex	4801	4	< 0.001
		+ Host plant	4814	4	ns

Table 2.3. Beak length differences in *Leptocoris tagalicus*. Linear mixed effect model including sex and host plant as fixed effects and sampled populations as random variable. The significance of the fixed factors was estimated using likelihood ratio tests between the Akaike Information Criteria (AIC) of the full model with those of the nested models excluding the appropriate variable (df, degrees of freedom).

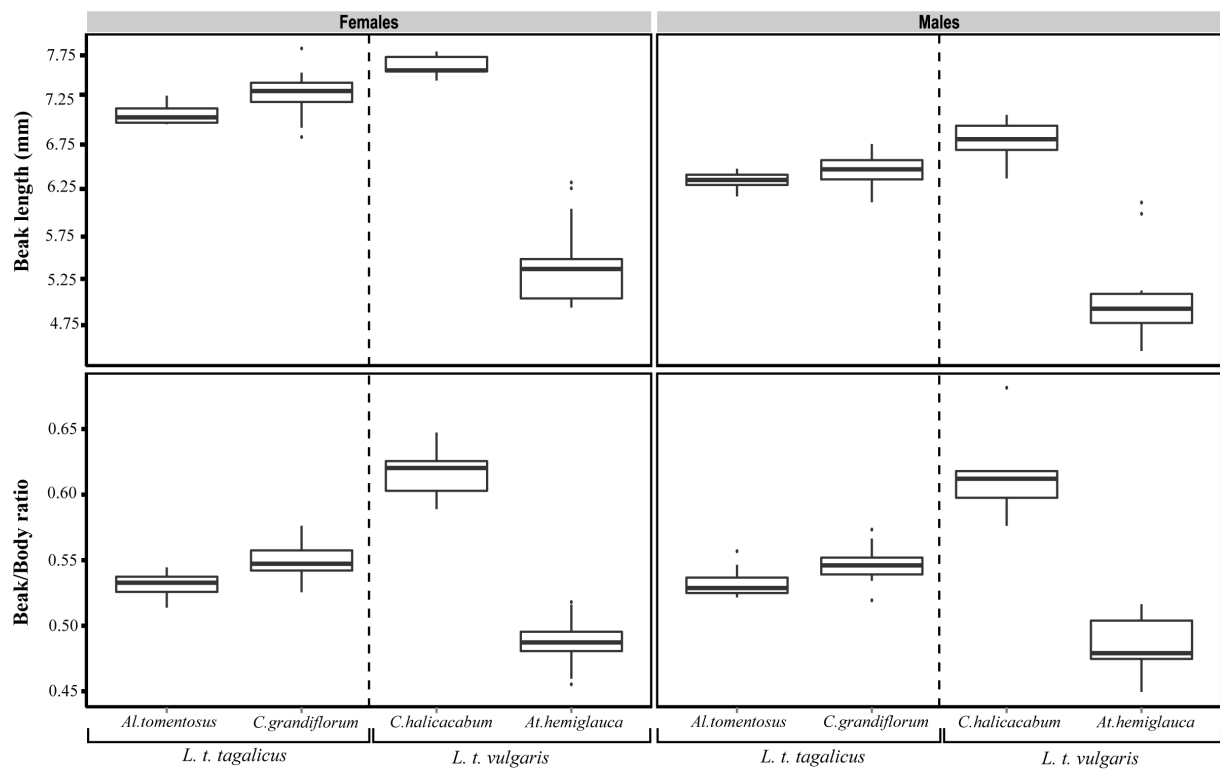


Figure 2.4. Box-plots of the grand means of morphological differences between both *Leptocoris* subspecies, and different host lineages. Boxes represent the first and third quartiles, whiskers represent 5 and 95% percentiles. Population means outside this range are represented as dots.

Nucleotide diversity (π) was relatively similar among lineages, ranging between 0.0037 and 0.0053. Bayesian and ML gene-tree reconstructions based on mtDNA were highly congruent and showed no geographic structure. Collapsing branches with support <0.50 resulted in one large polytomy (Fig. 2.5). Relationships between haplotypes were highly reticulated and showed no major breaks associated with *Leptocoris* subspecies, host plants, or populations (Fig. 2.6).

For the AFLP dataset both the g_1 statistic (-0.354 ; $P < 0.01$) and the PTP tests showed a significant phylogenetic signal (all permutations, $\alpha = 0.01$). However, the observed genetic differentiation is low and phylogenetic analyses (Fig. 2.7 A, B) showed only some evidence for structuring of the genotypes by subspecies: The MP tree (Fig. 2.7A) shows the *Halicacabum* bugs together with *vulgaris* and *tagalicus* in a single ladder-like, relatively well supported (75% bootstrap support). In contrast, the NJ tree (Fig. 2.7B) showed *vulgaris* and *tagalicus* as reciprocally monophyletic clusters but with only 70% posterior probability. The NJ tree also shows individuals collected on the introduced *C. halicacabum* as a fairly cohesive group (Fig. 2.7B). However, the support for this “cluster” is low (40% bootstrap support) as expected if these individuals represent a recently derived lineage.

2.3.4 Population structure

To infer the number of genetic demes we first used STRUCTURE. For our entire data set the estimated log probability [$\ln P(D)$] plateaus at $K = 3$ with the second-order rate of change of K (ΔK) reaching its maximum at $K = 2$, suggesting the existence of two (potentially three) distinct ancestry clusters. Assuming $K = 2$, the first cluster grouped all *tagalicus* specimens, while the second was representative of

the *vulgaris* form (Fig. 2.7C). There were 21 individuals exhibiting admixed ancestry ($q = 0.1-0.9$). Each of these individuals was initially classified as *vulgaris* and they were all collected in northern Australia, where the native *At. hemiglauca* co-occurs in sympatry with the introduced *C. halicacabum* (Fig. 2.7C). Whereas all individuals collected on *C. halicacabum* ($n=15$) were clearly admixed, only certain individuals on the native *At. hemiglauca* population showed some degree of admixture (Fig. 2.7C). Most of these admixed *vulgaris* bugs were found on the sympatric Victoria River population (Fig. 2.8).

Although STRUCTURE clustering with $K=3$ showed *vulgaris* bugs as admixed between (Fig. 2.4C), multivariate analyses (DAPC) showed that the BIC reached its minimum value at $K = 3$ suggesting that a subdivision into three clusters should also be considered. The first principal component explained 26% of the genetic variance and differentiated cluster 1 (*tagalicus*) from clusters 2 and 3 (*vulgaris*), while the second principal component explained 19% of the variance, displaying the difference between *vulgaris* populations on *At. hemiglauca* and *C. halicacabum* (Fig. 2.7D). The mean cluster membership probabilities based on the retained discriminant functions were ≥ 0.99 , and no individuals showed traces of admixture (*i.e.* $< 90\%$ membership in a single cluster). The two *Leptocoris* subspecies showed only modest genetic differentiation ($F_{ST} = 0.173$, $P < 0.001$). Within the *vulgaris* lineage there was relatively little differentiation between hosts ($F_{ST} = 0.106$, $P < 0.001$). All groups showed similar heterozygosity values and we found no significant differences between either subspecies ($H_{J-VULGARIS} = 0.102 \pm 0.004$; $H_{J-TAGALICUS} = 0.105 \pm 0.004$) or hosts ($H_{J-A.HEMIGLAUCA} = 0.106 \pm 0.004$; $H_{J-C.HALICACABUM} = 0.097 \pm 0.038$). The average proportion of AFLP loci showing significant pairwise linkage disequilibrium was

similarly low in each group: *vulgaris* (4.73%), *tagalicus* (4.22%) and *halicacabum* (4.50%) bugs.

2.3.5 Genome-wide admixture estimates and hybrid identification

Population structure analyses showed significant introgression in northern Australia. Detailed analyses of genome admixture revealed that in this region bugs feeding on *C. halicacabum* showed an intermediate assignment to each genetic cluster ($q_{\text{mean}} = 0.52$; $q_{\text{median}} = 0.48$); whereas bugs collected on *At. hemiglauca* showed very little introgression ($q_{\text{mean}} = 0.89$, $q_{\text{median}} = 0.88$; permutation-test $P < 0.001$; Fig. 2.8). Overall, we found no significant correlations between any of the morphological variables and the admixture index (Fig. 2.9). In the sympatric population of Victoria River the differences in admixture (q - values) between *vulgaris* and *Halicacabum* bugs were statistically significant ($q_{\text{VULGARIS-SYMPATRY}} = 0.90$, mean; $q_{\text{HALICACABUM-SYMPATRY}} = 0.63$, $P = 0.0024$), but less so than those observed between these bugs in allopatric populations ($q_{\text{VULGARIS-ALLOPATRY}} = 0.96$, $q_{\text{HALICACABUM-ALLOPATRY}} = 0.46$, $P < 0.001$). Analyses based on the hybrid index (H) of the individuals were almost identical, and there was a strong correlation between H - and q - values (Spearman- $\rho = 0.797$, $P < 0.001$) indicating our admixture estimates do not depend on pure parental samples being identified. The observed distribution of dominant genotypes at the 1,285 polymorphic AFLPs in the admixed balloon vine population was very different than that expected if this population was comprised solely by F_1 individuals (random permutation test $P = 0.0051$). While the *Halicacabum* bugs showed 6 fixed markers (*i.e.* AFLP band present) that are not found in any of the other populations, the number of private alleles in *tagalicus* and *vulgaris* was 39 and 29 respectively. Moreover, compared with the two putative ancestral lineages, the *Halicacabum* bugs

showed 508 loci (~ 40%) segregating at intermediate frequencies *tagalicus* and *vulgaris*.

2.4 Discussion

Introduced, naturalized plants are useful models for investigating the roles of selection and genetic variation in the formation of novel herbivorous insect lineages (Carroll et al. 2003, Feder et al. 2003). Intraspecific hybridization may be an important source of genetic variation for adaptation to potential new hosts because these plants may attract multiple, differentiated native insect populations, and trigger novel genotype by environment interactions that lead to host-shifts and ultimately, speciation (Schwarz *et al.* 2005). Here we suggest that the introduction of the annual balloon vine (*C. halicacabum*) to Australia 200-400 years ago has resulted in the rapid establishment of a new soapberry bug lineage characterized by a relatively small body size and a long beak suited to efficiently reach the seeds centered in the vine's large fruits. If the *Halicacabum* bugs have a hybrid origin their genome should be a blend of alleles derived from the two putative parental subspecies, *tagalicus* and *vulgaris*. In agreement with this prediction, all of our genetic analyses classified *Halicacabum* bugs as recombinants between these two lineages. Whereas all populations share mitochondrial lineages as if they were a single genetic group, Bayesian clustering (STRUCTURE) on nDNA revealed the existence of two distinct lineages, with *Halicacabum* bugs showing admixed identities. Consistent with this finding, hybrid-index analyses also showed evidence of mixed ancestry in *Halicacabum* bugs with most individuals (65%, n= 10) showing intermediate (0.4 - 0.5) hybrid indices between the two putative ancestors. Moreover, the *Halicacabum* bugs showed the highest number of loci segregating at intermediate frequencies and

fewer private alleles.

Alternatively, these patterns could arise if *L. tagalicus*, *L. vulgaris* and the Halicacabum bugs represent three distinct non-hybrid lineages derived from the same common ancestor species with the Halicacabum bugs harboring more ancestral polymorphism than the other two taxa and where ecological competition between *vulgaris* and Halicacabum bugs have lead to morphological character displacement. However, the inferred history of the annual balloon vine in Australia, the strict association between this lineage and the vine, and the lack of morphological differences between allopatric and sympatric populations of Halicacabum bugs are less consistent with this scenario.

To demonstrate that hybridization between these two lineages has resulted in a host-shift (*i.e.* the successful establishment of ecologically distinct lineage) one has to rule out the possibility –even if unlikely– that the Halicacabum bugs represent a “hybrid sink” of sterile individuals sustained by continuous immigration from the parental types. If so, the hybrid lineage should be composed of F₁ hybrids and would not be genetically differentiated from *vulgaris* and *tagalicus* (Gompert *et al.* 2006). Our results indicate that this not the case.

First, distribution of Halicacabum bug genotypes is very different from that expected in F₁ individuals. Second, the small fraction of AFLP-loci showing significant linkage disequilibrium in this lineage (~ 4%) is very similar to that

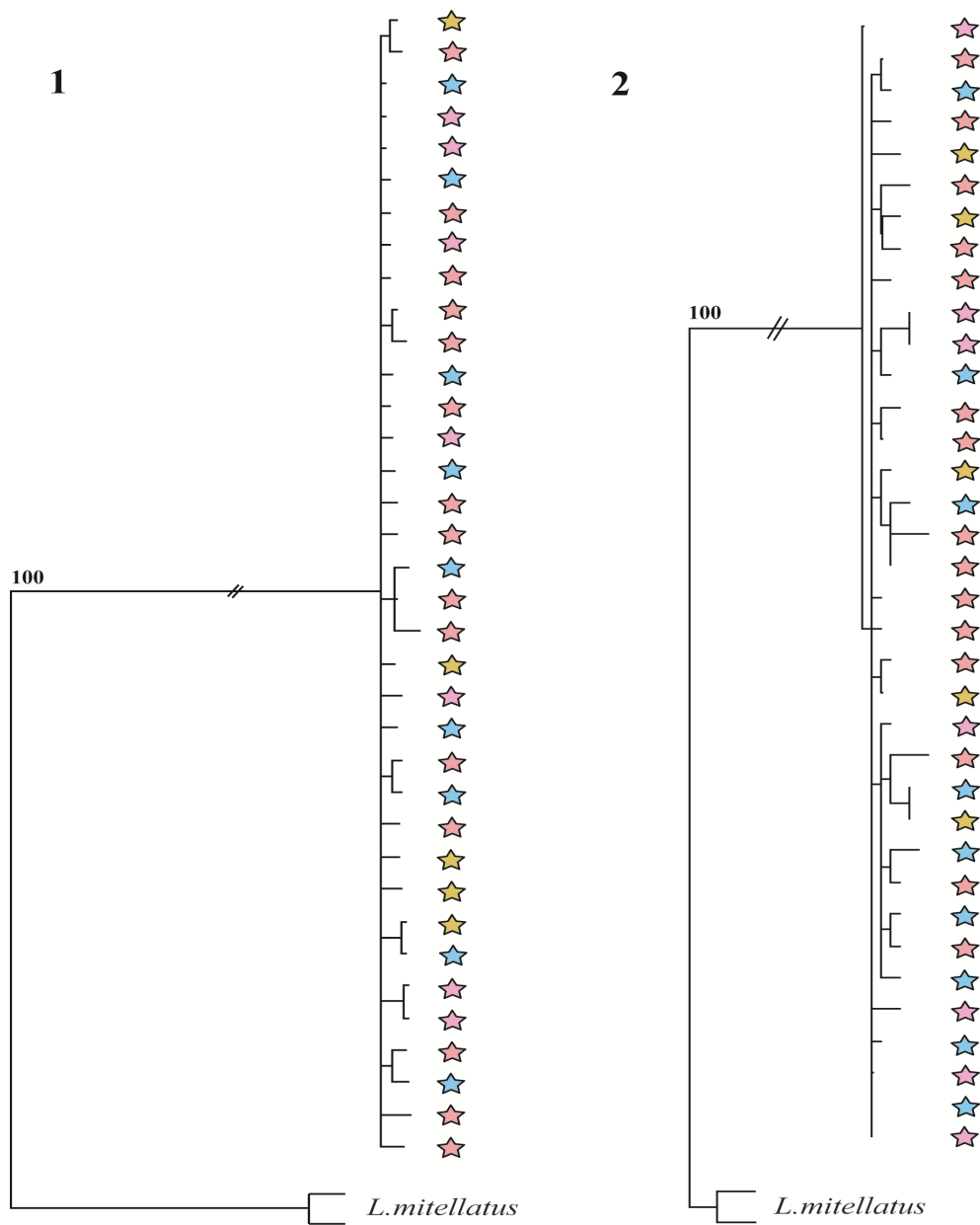


Figure 2.5. 1, Bayesian tree and 2, ML tree. Colors represent different host plants, Introduced: *C. halicacabum* (purple), *C. grandiflorum* (light blue). Native: *At. hemiglauca* (red), *Al. tomentosus*.

1



2

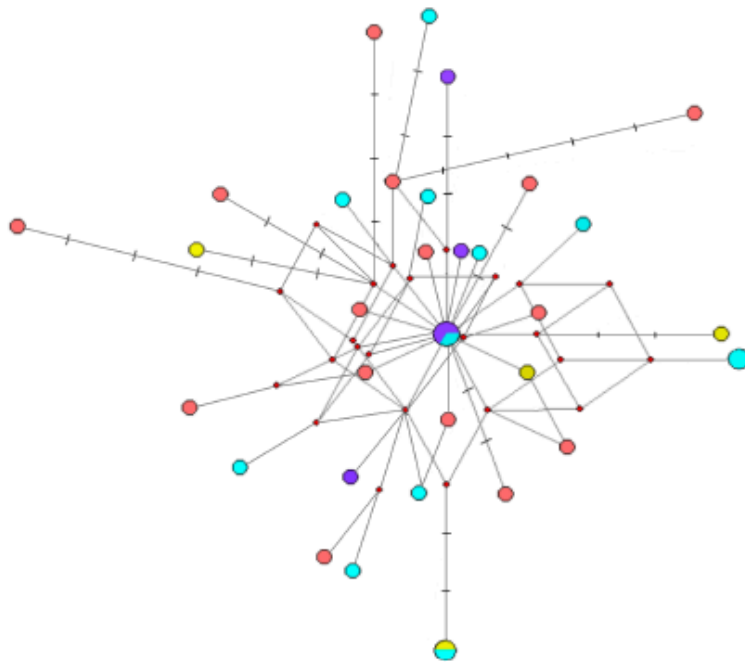


Figure 2.6. 1) Map of Australia showing the locations of the populations used in the genetic analyses. Colors represent different host plants. Native: *At. hemiglauca* (red) and *Al. tomentosus* (yellow); Introduced: *C. halicacabum* (purple) and *C. grandiorum* (light blue), while different shapes mark different populations within each host. 2) Median-joining network describing the evolutionary relationships between mtDNA haplotypes of *Leptocoris tagalicus*. Each haplotype is represented by a circle whose area reflects the overall number of copies observed and whose color-coding indicates the populations where the haplotype is found. Line length is proportional to the number of differences between haplotypes (tick marks). Small red circles represent non-sampled ancestral haplotypes that were reconstructed by the MJ algorithm.

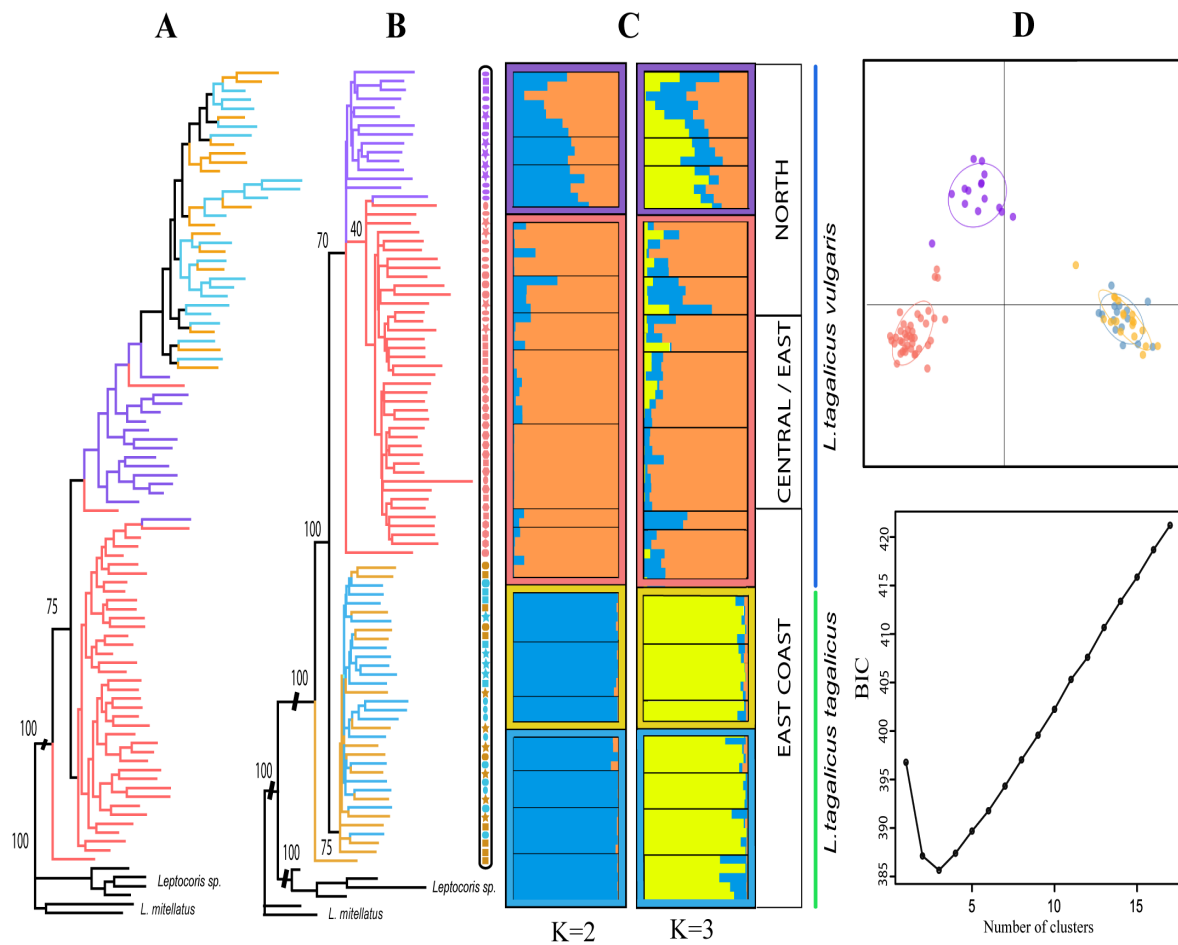


Figure 2.7. Phylogenetic structure of *L. tagalicus* based on ~1,200 EcoRI-MseI AFLPs. A. B. Maximum parsimony and Neighbor Joining trees. Numbers represent the bootstrap support (n=1,000). C. STRUCTURE analysis. The color of the individuals (rows) represents the proportion of their genome assigned to the K=2 and K=3 inferred clusters in the model-based admixture analyses D. Multivariate analyses (DAPC) Dots represent each individual. Groups are delineated by inertia ellipses. The analysis accounted for 60% of the genetic variation in the data set of which 26% and 19% are explained in first (x-axis) and second (y-axis) principal components. Colors represent different host plants. Native: *At. hemiglauca* (red) and *Al. tomentosus* (yellow); Introduced: *C. halicacabum* (purple) and *C. grandiflorum* (light blue), whilst different shapes mark different populations within each host.

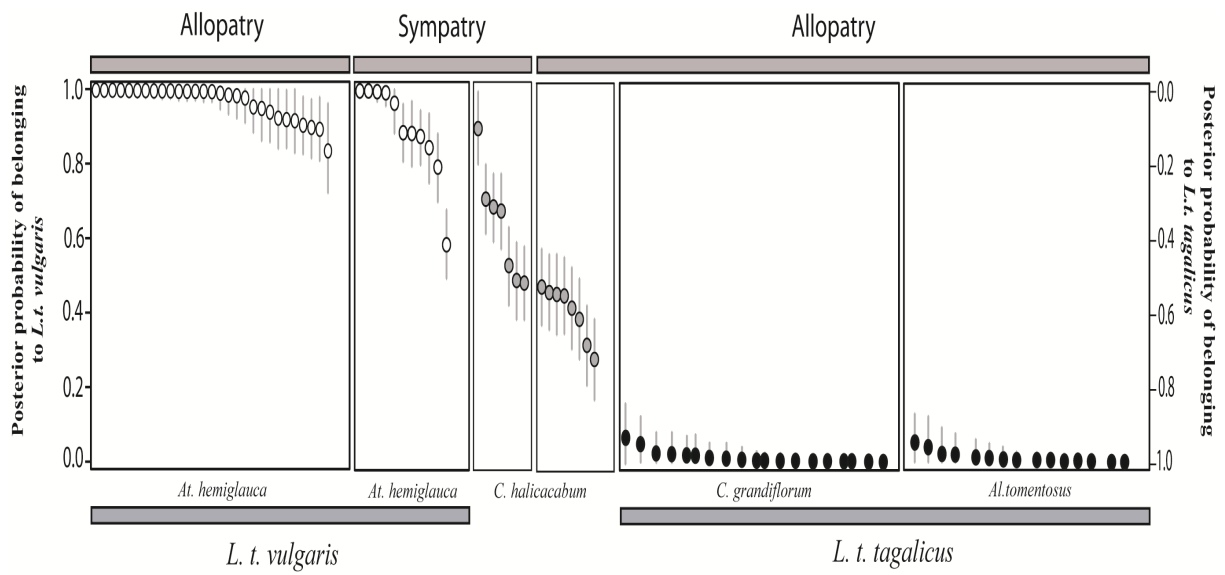


Figure 2.8. Assignment of ancestry of the annual vine bugs from two parental populations (grey dots). Black and white dots represent *tagalicus* and *vulgaris* individuals respectively. *vulgaris* individuals are further split into allopatric and sympatric populations. In sympatry, two different hosts, the non-native annual vine (*C. halicacabum*) and native whitewood tree (*At. hemiglauca*) coexist sometimes only a few meters apart.

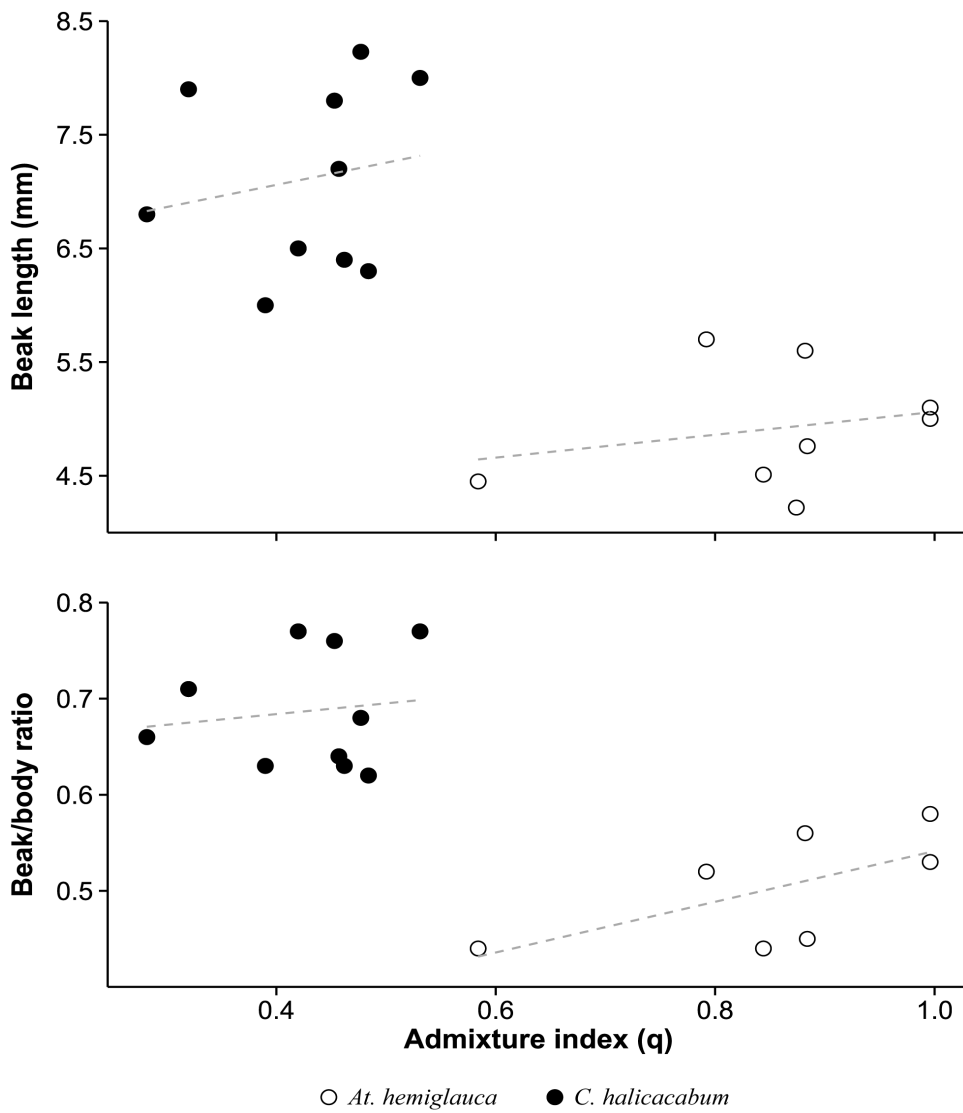


Figure 2.9. Relationships between admixture index, beak length and beak/body ratio. Black and white dots represent individuals collected on annual vine (*C. halicacabum*) and whitewood (*At. hemiglauca*) respectively.

observed in the two putative parental subspecies. Third, multivariate analyses (DAPC) indicate a subdivision into three clusters: ‘pure’ *vulgaris*, ‘pure’ *tagalicus*, and the hybrid populations, with the *Halicacabum* bug lineage showing six fixed unique alleles that are not present in either of the parental species. Thus, the most parsimonious explanation for our results is that hybridization of small *vulgaris* bugs with long-beaked *tagalicus* bugs resulted in bugs with a phenotype (long beaks and high beak/body ratios) that enable the exploitation of the inflated fruits of the annual balloon vine, allowing the establishment of a new recombinant lineage (*sensu* (Arnold 1997; Rieseberg 1997)).

Hybrid host shifts represent an ecologically robust scenario for animal hybrid speciation because they offer a potential mechanism for reproductive isolation through differential adaptation to a new ecological niche (e.g. (Emelianov *et al.* 2003; Gompert *et al.* 2006; Kuusela *et al.* 2007; Mercader *et al.* 2009; Schwarz *et al.* 2005)). For example, hybridization is thought to underlie the recent ‘instantaneous’ speciation in the native North American *Rhagoletis* fruit flies breeding on naturalized hybrid East Asian honeysuckles (*Lonicera*) in the eastern United States (Schwarz *et al.* 2005). *Lonicera* flies are reproductively isolated from other flies and show a unique allelic mixture of two native fly taxa. Thus, a new stable lineage has formed without a change in chromosomal number as the result of interbreeding between diverged parental lineages (homoploid hybrid speciation). A similar process has been reported for the origin of alpine-adapted *Lycaeides* butterflies (Gompert *et al.* 2006).

Do the *Halicacabum* bugs represent a case of hybrid speciation driven by the introduction of a new host? When a host shift occurs a range of outcomes is possible

and demonstrating the existence of reproductive barriers between the bug population on *C. halicacabum* and its putative parentals is, therefore, a crucial next step to address this question (Rieseberg 1997). Though limited, our current data from Victoria River, NT (where the introduced *C. halicacabum* and the native *At. hemiglauca* grow only a few meters apart) are consistent with the existence of ongoing, but reduced, gene flow between the hybrids and one of the parental species. In this population the beaks of the Halicacabum and *tagalicus* bugs are as distinct as bugs found in allopatric populations. However, though still distinct from *vulgaris*, the hybrid indexes of Halicacabum bugs were significantly higher than in allopatry, suggesting that backcrossing occurs in sympatry. Further field and laboratory studies are needed to determine the degree of isolation of Halicacabum bugs. Genome-wide studies using co-dominant markers should be implemented and laboratory crosses should be carried out to try to reconstruct the phenotype of the Halicacabum bugs. In addition, the nature of the reproductive barriers between the hybrids and the putative parental species should be carefully determined using a combination of host preference, mate choice and host-specific fitness experiments.

In contrast to the bug populations using *C. halicacabum*, the bugs feeding on the more recently introduced *C. grandiflorum* showed no evidence of admixture, and they are genetically indistinguishable from the *tagalicus* populations found on the native *Al. tomentosus*. However, bugs feeding on the *C. grandiflorum* have longer beaks than those feeding on *Al. tomentosus*. Cross-rearing experiments have shown that this host-associated difference in beak length is heritable and that it is not just related to phenotypic plasticity (Carroll *et al.* 2005a). Thus, it seems that host-specific adaptation has preceded the development of population structure. A more detailed

evaluation of the genetic variation is needed to assess the relative roles of gene flow and shared ancestral polymorphism to the lack of genomic differentiation between the bugs feeding on these two hosts. Similarly, in North American *Jadera* soapberry bugs for which *Cardiospermum* is the native host, rapid diversification on introduced Asian host plants is based on previously rare or formerly unexpressed interactions among existing genes, suggesting that that shift to perennial vine may be strongly influenced by epistatic and dominance interactions. (e.g. (Carroll 2008; Dingle *et al.* 2009)). More genetic studies are needed to describe the factors controlling host adaptation in this branch of the Australian system, and for practical purposes to evaluate the potential for *L. t. tagalicus* to evolve enhanced biological control capacity against invasive perennial balloon vine. We predict that further results will confirm the importance of hybridization and epistasis as two main mechanisms that facilitate the adaptation of plant-associated insects to new hosts.

2.5 Conclusions

Haplotype network did not show any major association with either *Leptocoris* subspecies or host plants, and no association with the geographic locations. Taken altogether, the most parsimonious explanation for the observed genetic, morphological and distributional modeling patterns is that hybridization of small *vulgaris* bugs with long-beaked *tagalicus* ones resulted in bugs with a phenotype (long beaks and high beak length/body length ratios) that allowed the offspring to exploit the inflated fruits of the annual balloon vines and freed the new lineage from parental species competition, providing the opportunity for ecological isolation to arise.

CHAPTER 3 - GENETICS OF HOST-SHIFT AND ADAPTATION IN *JADERA* BUGS OF FLORIDA, USING CANDIDATE GENE APPROACH.

3.1 Introduction

Plant species vary inter-and intra-specifically in numerous morphological, phenological and chemical traits and insect's responses to plant trait variation are responsible for the rapid diversification and high diversity of phytophagous insects (Agrawal *et al.* 2012; Groman & Pellmyr 2000; Matsubayashi *et al.* 2010; Thomas *et al.* 2003; Wood & Keese 1990). Although conceptually this does not present any challenge, the ecological and genetic mechanisms behind host shifts are yet to be understood.

From a genetic point of view, the adaptation of phytophagous insects to novel hosts might occur through new mutations, existing genetic variation, hybridization, and the introgression of pre-adapted genetic elements that originated outside of currently adapting populations (Feder *et al.* 2003b). Understanding the role of each of these genetic mechanisms in host shifts requires identifying the genes responding to host-related selection. In this chapter, I use a candidate gene approach to identify genes that may be associated with the anthropogenic host shift in the soapberry bug, *Jadera haematoloma*. Florida populations of this species feed on either native balloon vine (*Cardiospermum corindum*) or an introduced golden rain tree (*Koelreuteria elegans*), commonly planted since the 1950s (Carroll & Boyd 1992). Morphological, nutritional and ecological differences in host plants appear to have driven the evolution of soapberry bugs (Carroll *et al.* 2003; Carroll *et al.* 1997; Carroll *et al.* 1998). Morphologically, while adult *J. haematoloma* currently feeding on balloon vine closely resemble the museum specimens collected prior to the introduction of the introduced

tree, those now living on the golden rain tree have diverged beyond their ancestral phenotypic range, and currently show beaks that are significantly shorter (Carroll & Boyd 1992). Golden rain-bugs also mature faster (~25% decrease), have higher fecundity (~ 2 times increase), and survive better (~40% increase) in the introduced host than the balloon vine ones (Carroll *et al.* 1997; Carroll *et al.* 1998). As a result, these host-specific populations have been considered different ecomorphs (Carroll *et al.* 2001; Carroll *et al.* 1998).

Classic quantitative genetic studies and selection experiments have shown that the size of the beak in *Jadera* is under genetic control (Carroll *et al.* 2005a; Dingle *et al.* 2009) and at least one QTL analysis suggests that this trait is controlled by a few loci of relatively large effects (Yu & Andres 2014). Thus, in this chapter, I target three candidate developmental genes (*Distal-less (Dll)*, *dachshund (dac)* and *homothorax (hth)*) that are known to play a potential role in the insect's mouthpart and appendage development in arthropods (Angelini & Kaufman 2004; Panganiban *et al.* 1997; Prpic & Tautz 2003). The first candidate gene, *Distal-less (Dll)*, encodes a homeodomain transcription factor, expressed in distal structures of appendages during development, their expression and function is widely conserved in all invertebrate appendages (Cohen *et al.* 1993; Ishimaru *et al.* 2015; Panganiban *et al.* 1997). *Dll* is required in the labium for development of the most distal portion (Angelini & Kaufman 2004, 2005). The second one, *dachshund (dac)* is a transcription factor, that is vital for proper differentiation of a subset of segments in the developing leg, and has a function in sensory structures including the eyes (Mardon *et al.* 1994; Yang *et al.* 2009). *Dac* is expressed throughout the length of the embryonic mandibular, maxillary limb buds and of proximal domain in the labium (Angelini & Kaufman 2004).

Finally, the third candidate gene, *homothorax* (*hth*) encodes a homeodomain transcription factor. Expression of *hth* is essential for proper elongation of the stylets; it is also expressed in proximal portion of the labium, throughout the length of the mandibular and maxillary limb buds (Angelini & Kaufman 2004, 2005; Angelini *et al.* 2012).

Host specialization is obviously not only related to morphological changes but many more phenotypic changes usually accompany host specialization. Amongst them, changes in host preference and host toxins have been particularly well documented in other insects (Aldrich *et al.* 1990; Bramer *et al.* 2015; Dalla *et al.* 2013; Djamgoz *et al.* 1998; Dobler *et al.* 2012; McBride 2007). Different plant lineages have evolved independently to produce a vast variety of highly efficient toxins to avoid herbivores (Dalla *et al.* 2013). The seeds of both the *Cardiospermum* vines and the golden rain tree contain cyanolipids, alkaloids, saponins, glycosides, tannins, flavonoids and cardiac glycosides (Aldrich *et al.* 1990; Patil *et al.* 2010; Raza *et al.* 2013; Suresh *et al.* 2012).

Cardiac glycosides are cardenolides and are known to inhibit the Na⁺/K⁺-ATPase pump, a heterodimeric member of P-type ATPases (Hansen 1984; Yoda & Yoda 1982). Although the α -subunit of this enzyme is highly conserved (Dalla *et al.* 2013; Dobler *et al.* 2012; Horisberger 2004), phytophagous insects feeding on hosts producing cardenolides show characteristic amino acid substitutions conferring resistance to these toxins (Agrawal *et al.* 2012; Dobler *et al.* 2012; Petschenka *et al.* 2013; Zhen *et al.* 2012). Thus, it is likely that *Jadera* bugs also show “resistant” ATP-ase alleles, which may be different than those, described in other heteropterans.

To test this hypothesis in this chapter I have characterized the genetic variation found in the ATPase gene in Florida populations of *Jadera haematoloma* feeding on either host.

Host preferences in phytophagous insects are controlled by chemosensory systems that detect volatile cues (*i.e.* odorants) and allow them to detect food, predators, mates and oviposition sites (Andersson *et al.* 2011; Deisig *et al.* 2010; Hua *et al.* 2012; Rospars *et al.* 2014). Insects sense these odorants with specialized odorant receptors (ORs) expressed in the antennal receptor neurons. ORs bind both odorants and a co-receptor called *Orco* (Odorant Receptor-Coreceptor), which in turn activates the ion channels of the olfactory neurons (see Fig. 3.1) (Chen & Luetje 2012; Cooper *et al.* 2000; Leal 2013; Palczewski *et al.* 2000; Sato *et al.* 2008).

Odorants vary dramatically across different host species (Hallem *et al.* 2006).

Consequently, ORs genes are fast evolving genes (Hua *et al.* 2012) and the number of OR genes vary across related species of phytophagous insects. Because activated olfactory receptors directly interact with the *Orco* aminoacid substitutions in ORs may require matching changes in the *orco* gene, even if some of the *orco* domains are known to be highly conserved across all insect taxa (Jones *et al.* 2005; Taylor *et al.* 2012). In the final part of this chapter I test to see if the *orco* gene shows any differences between the two soapberry bug ecomorphs found in Florida.

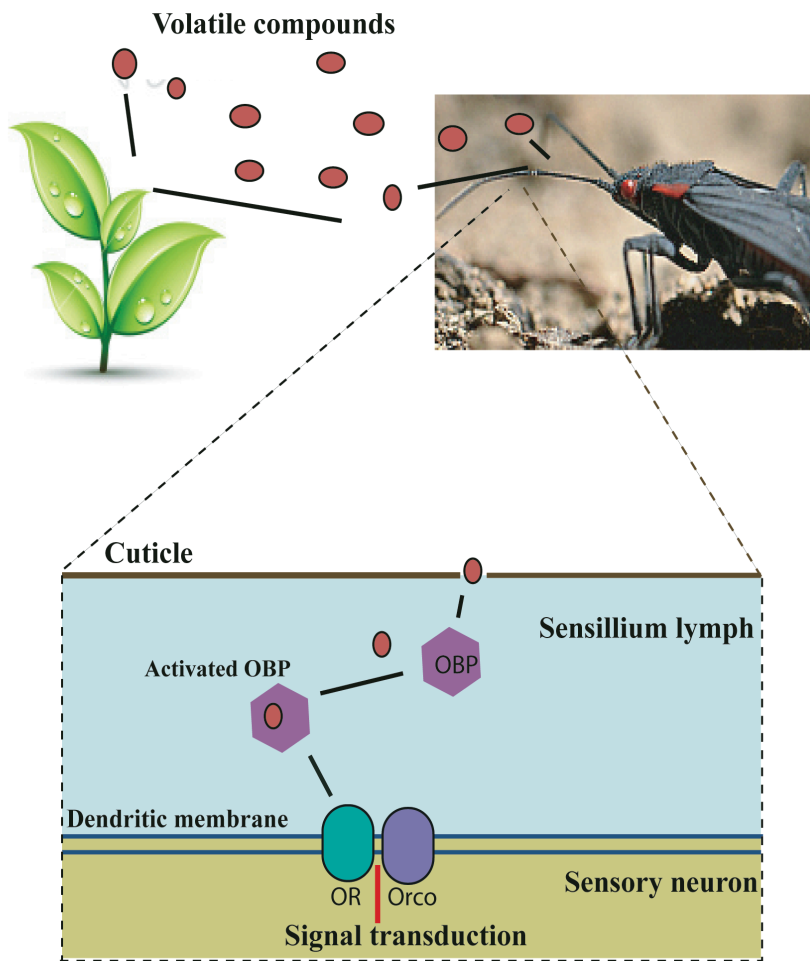


Figure 3.1. Schematic representation of activation of ion channel by volatile compounds (Odorants) produced by plants, Interaction between the ORs and *orco* receptors.

3.2 Materials and Methods

3.2.1 Characterization of the candidate genes

All of the three developmental genes (*dac*, *Dll*, *Hth*), *orco* and Na⁺/K⁺-ATPase are expressed in their appendages and throughout the life cycle of the bugs. So the antennae and beak from adult male and female *Jadera* bugs were harvested using forceps and stored at -80°C. Total RNA was extracted from the frozen tissues using Trizol, according to the manufacturer's instructions. First strand cDNA was synthesized from approximately 1 µg of total RNA using ThermoScript reverse transcriptase (Invitrogen, Carlsbad, CA) and random hexamers using conditions recommended by the manufacturer.

orco gene was amplified using degenerate primers designed to conserved amino acid stretches (IKAWYPW, AIKYWV and VCQQCQK) of previously identified *orco* sequences. Amplifications were performed using 0.7 µL of the cDNA template and 5 µL of each primer with platinum-Taq DNA polymerase and thermo cycler conditions as used in Hull et al., 2012. To amplify the Na⁺/K⁺-ATPase (α -subunit), *Dll*, *dac*, and *hth*, degenerate primers were designed to highly conserved regions from the multiple sequence alignment built from previously available annotated insect sequences (see Table 3.1). Using the small fragments as a template, genome walk was performed to cover the N-terminal and C-terminal to cover all the main domains for all the genes.

No	Primer Name	Oligo sequence
1	orco_newF1 orco_newR1	5'-GCGACCTCTGGCCGAACATC-3' 5'-GCACCAAGCACAGAAGCAAAAAGA-3'
2	Dll_newF1 Dll_newR1	5'-ACCACCTAAAGATGAGAAATG-3' 5'-GATGATGATGGGGATGGTGGTGAG-3'
3	dac_PF2 dac_PR2	5'-TCCCAGCAGATTAAGAAGCACAGG-3' 5'-GCGGCAGGCGACGACACG-3'
4	hth_LTF1 hth_LTR1	5'-TTTGAAAAGTGCGAGTTGGCTACG-3' 5'-TTGTTAATCCGGTGTCTTGTGCTA-3'
5	fATPa rATPa NT_atpF1 NT_atpR2 CT_atpR2 rATPc	5'-ATGACNGTNGCNCATATGTGGT-3' 5'-ATNGGGTGGTCNCCNGTNACCAT-3' 5'-TRYTVTGGGTTGGYGCHAT-3' 5'-CTTTGAATCCAGGGCTTGTT-3' 5'-TCTTATGTCTATGATTGATCCACC-3' 5'-ATNGCNGGNACCATGTCNGTNCC-3'

Table 3.1. List of Primers

	Population Site	Host
1	Homestead Ft	<i>Cardiospermum corindum</i>
2	Key Largo	<i>Cardiospermum corindum</i>
3	Gainesville	<i>Koelreuteria elegans</i>
4	Orange City	<i>Koelreuteria elegans</i>
5	Ft. Myers	<i>Koelreuteria elegans</i>
6	Homestead Ft	<i>Koelreuteria elegans</i>

Table 3.2. List of *Jadera* populations from Florida and their associated hosts.

3.2.2 Identification of genetic variation among populations

Total RNAs were extracted from heads of 20 adult *Jadera* bugs from **two native** *C. corindum* populations (Homestead, Key Largo), **four derived** *K. elegans* populations (Gainesville, Orange City, Ft. Myers, and Homestead) (see Table 3.2, Fig. 3.2). cDNAs were synthesized as per manufacturer's instructions. The genes were amplified using the same PCR conditions used for their characterization. Purified (Exo/SAP) PCR products were sequenced on 3500-XL Genetic Analyzer using ABI 3.1 Big Dye Terminator chemistry.

3.2.3 Phylogenetic analyses

The populations of bugs living on different hosts are believed to evolve independent to each other, and so the mutations will be common among the population living on respective host plants. We employed multiple sequence analysis and phylogenetic analyses to identify the genetic differences between and within populations. The nucleotide sequences from populations (Table 3.1) for the genes *Dll*, *dac*, *orco* and fragment of Na⁺/K⁺-ATPase (α -subunit) were aligned using the program using MAFFT (Kato & Standley 2013). Phylogenetic relationships were estimated using two different methods 1) Neighbor Joining (NJ) and 2) Maximum Likelihood (ML), using 1000 bootstrap. If there are genetic differences associated with the host, then the populations feeding on different host-plants are expected to form monophyletic groups.

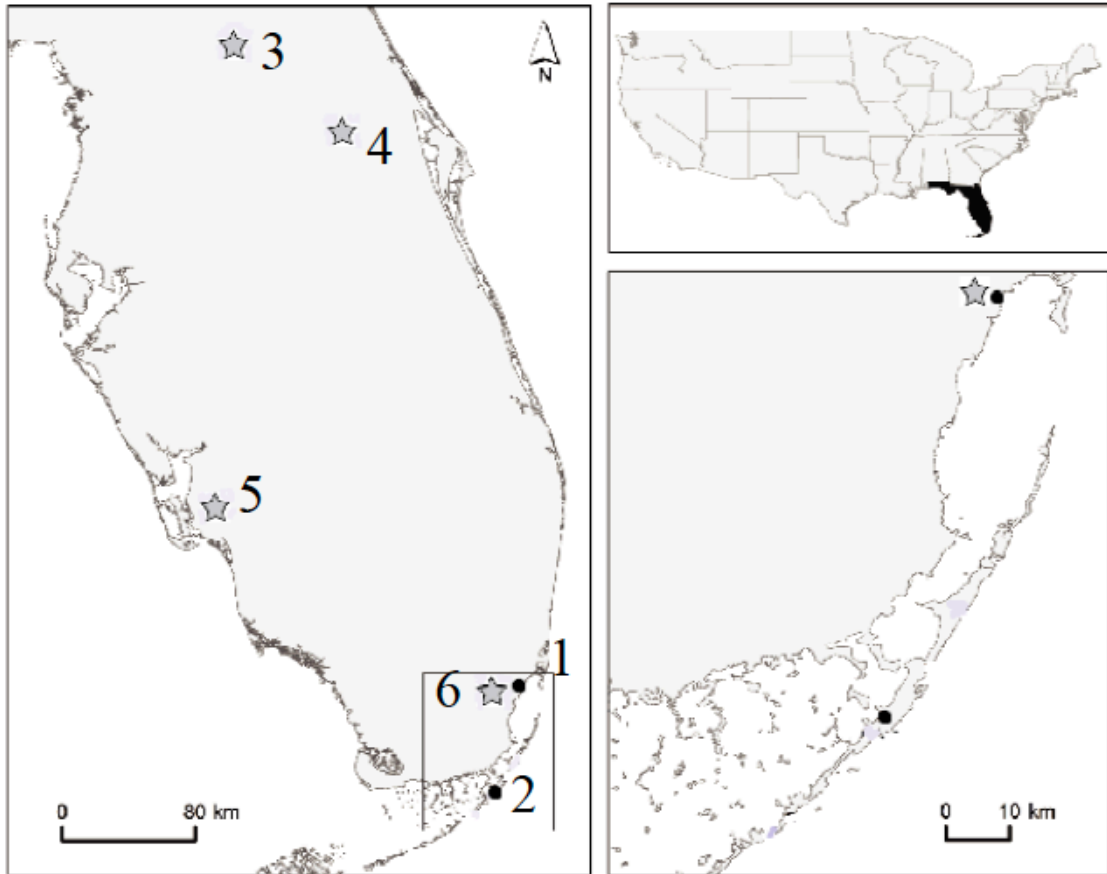


Figure 3.2. *Jadera haematoloma* populations and sampling sites. Upper right: map of the United States of America with Florida highlighted. Left: Florida. Lower right: close-up of the Florida Keys and part of the tip of the peninsula. Black circles represent known populations of the ancestral long-beaked ecomorph feeding on the native balloon vine (*C. corindum*). Stars represent known populations of the derived short-beaked ecomorph feeding on the introduced golden rain tree (*K. elegans*).

3.2.4 Population genetic analyses

Genetic diversity within and between populations was estimated by computing haplotype diversity (H) and nucleotide diversity (π). To analyze demographic events such as past population expansion and selection driving forces, we used statistical analysis Tajima's D (Tajima 1989), which uses the frequency of segregating nucleotide sites. Negative values of these tests are expected in populations that have undergone a recent population expansion, because expected alleles are more frequent than rare alleles; whereas positive values occur if rare alleles are eliminated from population following genetic bottlenecks (Tajima 1989). The analyses were implemented in the program DNASP (Librado & Rozas 2009).

The same DNASP (Librado & Rozas 2009) package was used to calculate the rate of amino acid replacement substitution (dN) relative to the rate of silent substitution (dS), $\omega = (dN/dS)$. For any set of amino acid residues, when $dN/dS = \omega = 1$, a neutral model of evolution cannot be rejected, whereas $\omega < 1$ indicates purifying selection and $\omega > 1$ indicates positive selection. The difference in these fixation rates provides an estimate of the selection pressures.

3.3 RESULTS AND DISCUSSION

3.3.1 Sequence analyses

We have sequenced *orco* (1400bps), and Na^+/K^+ -ATPase (2085bps), *Dll* (705bps), *dac* (1085bps), and *hth* (913bps). Sequence analysis showed that the genes are highly conserved, and the nucleotide diversities (π) were relatively low. Overall only 9 parsimonious informative sites were found among all the genes, and all the mutations were associated with individuals in the population rather than common for

a specific host-associated or population-associated. Further, population genetics analyses of *hth*, *dac* and Na⁺/K⁺-ATPase revealed high diversity of haplotypes, constituting ~35% of the total sample size. The combination of high haplotype diversity and low nucleotide diversity that we observed in our data may be representing a signature of a rapid demographic expansion from a small effective population size (Avice 2000; de Jong *et al.* 2011). Statistical tests these days are developed to test selective neutrality of mutations, and are implemented to detect such population growth (de Jong *et al.* 2011; Ramos-Onsins & Rozas 2002). The results of Tajima's D test were negative for all populations for all the genes (*hth* = -1.6302, *dac* = -1.7406, *Dll* = -1.2331, Na⁺/K⁺-ATPase = -0.6934 and *orco* = -1.18951), indicating a larger number of rare nucleotide site variants compared to the expected under a neutral model of evolution. The genes that are evolving under directional selection (i.e. positive or negative selection), their variants will be at a low frequency so nucleotide diversity will be low relative and will produce negative values. But, Tajima's D will also be affected by demography, because population expansion can also lead to an excess of rare alleles. So in this particular case, the results may be implying a case of recent population expansion (Navarro & Barton 2002; Roux *et al.* 2013). The phylogenetic analysis also did not reveal any association of population with any host (see Fig. 3.3)

It is very evident that *orco* interacts with most of the ORs and these ORs are fast evolving genes (Benton *et al.* 2006; Chen & Luetje 2012), but still our results showed very low nucleotide diversity ($\pi = 0.00025$) and low substitution rates (one synonymous and one non-synonymous). This shows that there is strong selection acting on these genes preventing any changes in the amino acid substitutions, and

may be it is very important for this gene to maintain the genetic architecture to hold specific structural conformation related to ORs interactions.

Jadera populations from Homestead and Key Largo were feeding constantly on *Cardiospermum* seeds, which have known cardenolide content. And the insect taxa that are feeding on these cardiac glycosides has amino acids substitution in their Na⁺/K⁺-ATPase α -subunit, helping the phytophagous insects to adapt and colonize the cardenolide containing host plants. Interestingly, our experiment revealed that the *Jadera* populations feeding on *Cardiospermum* do not have any substitutions or changes in their Na⁺/K⁺-ATPase α -subunit sequences. Instead, all the sequences are highly conserved. The α and β -subunits are known to interact with each other, mutations of different amino acids resulted in heat-sensitive enzymes, lower affinity for K⁺ and inactive protein (Becker *et al.* 2004). So highly conserved stretches of Na⁺/K⁺-ATPase sequences among the insect taxa represent the importance of maintaining a particular amino acid at a particular site. Multiple sequence alignment of *Jadera* Na⁺/K⁺-ATPase (α -subunit) with other insect taxa that live on hosts with and without cardenolides (see Fig. 3.4) have shown *Jadera* sequence clearly clumps with a bunch of other taxa that live on hosts plants free of cardenolides.

Based on the experiments conducted in Angelini's lab it is very obvious that the three candidate developmental genes we chose have a role in appendage developments in soapberry bugs and also in related *Oncopeltus fasciatus* (Angelini & Kaufman 2004). Our results have shown all three *dac*, *Dll*, *hth* genes are high conserved. For the gene *Dll*, the number of polymorphic sites is two and nucleotide diversity (π) = 0.00079, for the gene *dac* polymorphic sites is four and nucleotide diversity (π) = 0.00116. For the gene *hth*, the number of polymorphic sites is five (all

singletons) and nucleotide diversity (π) = 0.00165. Interestingly, *Dll* gene has zero non-synonymous substitutions to two synonymous substitutions. The gene *hth* gene has four non-synonymous substitutions and one synonymous substitution, but no changes are related to populations or host plants. Meanwhile *dac* gene has three non-synonymous substitutions to two synonymous substitutions.

3.4 Conclusions

The most possible parsimonious explanation for low genetic variability among these three developmental genes are 1) The interactions between the developmental genes may be playing a major role in pattern formation of appendages, so the changes in phenotype may be due to the differences in expression level or because of the epistatic effect. Also its been published that the host based differences in beak length and other performance measures in *Jadera* are largely the product of epistatic differentiation (Carroll 2007a; Carroll *et al.* 2003; Carroll *et al.* 2001). 2) Developmental genes are mostly pleiotropic, affecting several independent traits; indeed, many appendage-patterning genes likely to influence the beak length are known to affect other appendages. So the phenotypic changes that we observed may be rooting from a different set of genes.

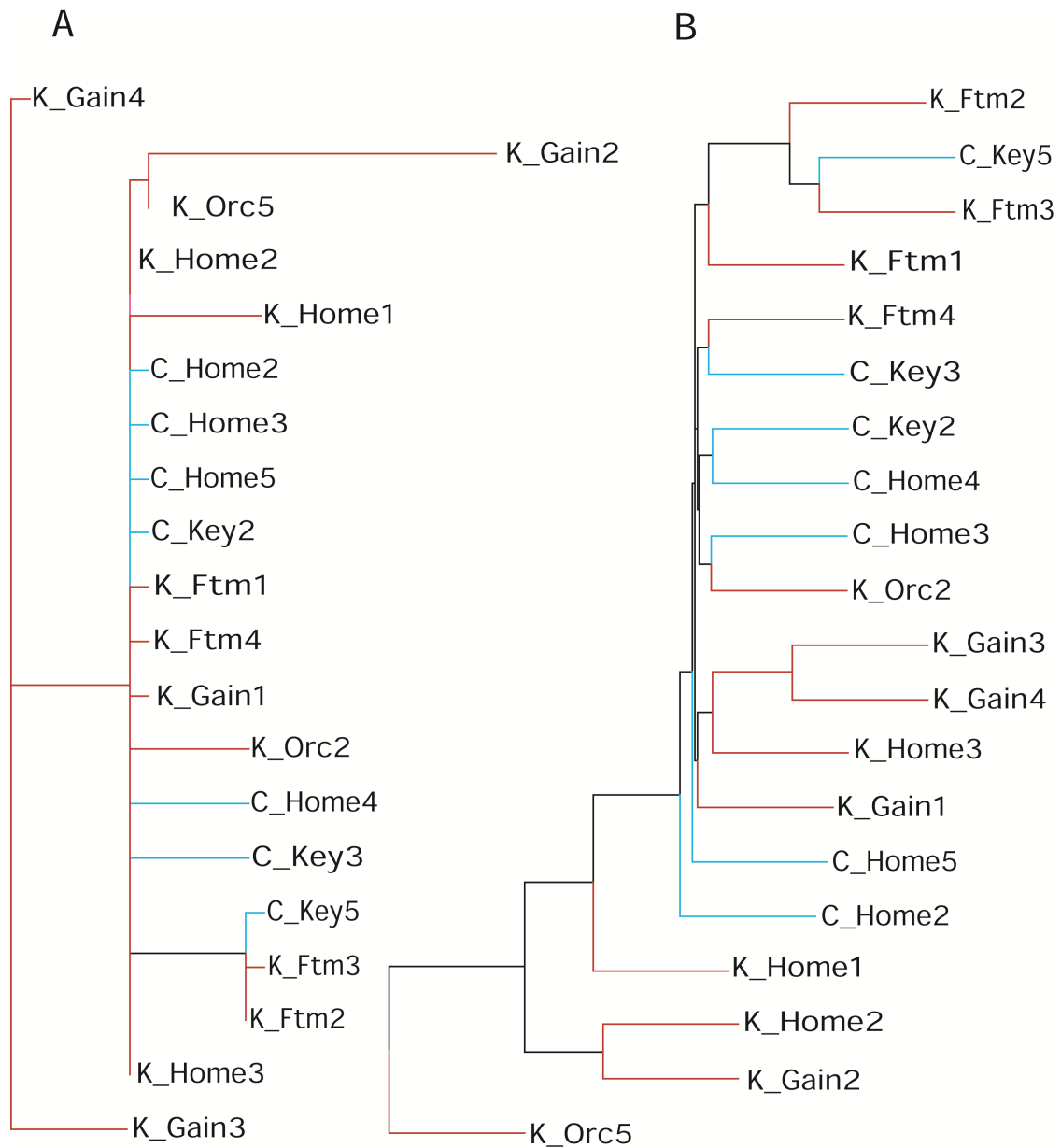


Figure 3.3. Phylogenetic trees with concatenated 2664bps from genes. A- Maximum Likelihood tree, B- Neighbor Joining tree. Red color represents *K. elegans* population and blue represents *Cardiospermum* populations. Home- Homestead, key-Key Largo Gain-Gainesville, Leesburg, Ftm- Ft. Myers.

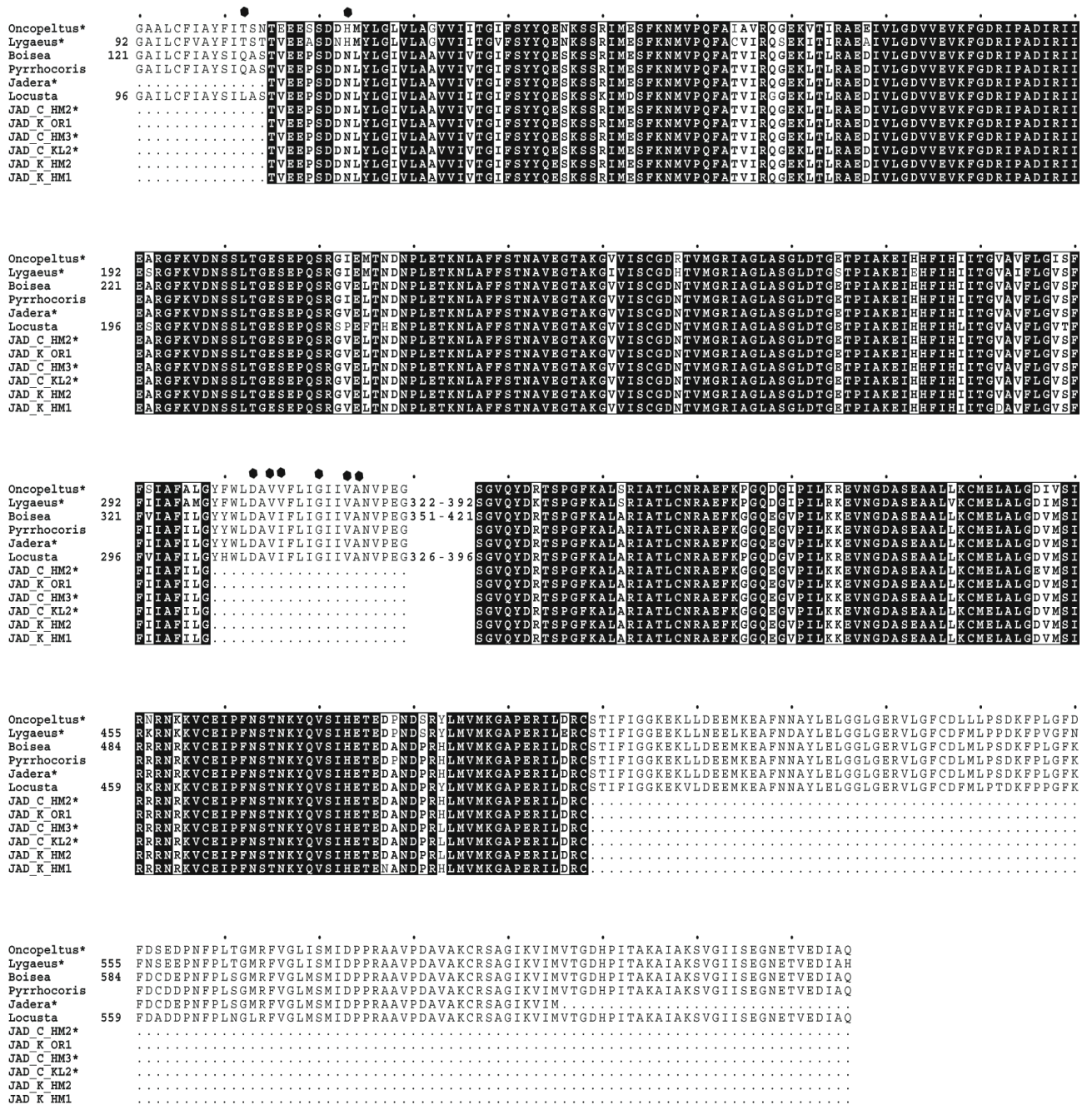


Figure 3.4. Multiple sequence alignment of Na^+/K^+ -ATPase (α -subunit) of representative sequences from different insect taxa. * Represents sequence from insects feeding on cardenolide-containing seeds. *Jadera** represents the gene fragment sequenced from a RNA pool of bugs populations feeding on cardenolide-containing host plant. Full-length sequences are given with their total number of amino acid residue numbers. Small hexagon symbols on the top of alignment block represents the amino acid substitutions involved in cardenolide binding.

CHAPTER 4 - CONCLUSIONS AND FUTURE DIRECTIONS

Soapberry bugs provide us with a unique opportunity to study the genetic mechanisms underlying host-shifts. The colonization of invasive balloon vines by Australian soapberry bugs (*Leptocoris spp.*; Chapter 2) greatly illustrates that host-shifts are not always facilitated by new mutations or standing genetic variation but by hybridization and adaptive introgression. Sets of genes that are candidates for host-shifts can be identified through various genetic tools, including selection experiments that test for correlated responses in candidates, QTL mapping, and comparisons of “host races” that test for specific genetic associations (Genome wide association studies (GWAS)). Most of these techniques lead to the identification of candidate sets of loci rather than specific alleles. Nevertheless, the identification of specific alleles that are involved in the adaptation to a new hosts plant is still necessary if we want to fully understand the molecular bases of host-plant colonizations. A potential way to identify potential adaptive alleles is to perform a candidate gene approach similar to that performed in Chapter 3. While my efforts during my Masters focused on structural changes in the candidate genes themselves it is clear that regulatory changes as well as gene duplication and gene loss, are likely to have a role in adaptive shifts. Future efforts should concentrate not only on the identification and characterization of other candidate genes (e.g. Odorant Receptors) but also on the study of the differential gene expression between soapberry ectomorphs feeding in different hosts. For example, recent advances in sequencing technology (NGS) allow us to generate reference transcriptomes. Then, an RNA-seq approach can be used to detect if the different ecomorphs have genetically diverged in their gene expression response to different host plants. Each of the methods used for identifying potential loci involved

in host adaptation has advantages and disadvantages; it is only through an integrative approach that we will unravel the genetic changes associated with the colonization of new hosts.

REFERENCES

- Agrawal AA, Petschenka G, Bingham RA, Weber MG, Rasmann S (2012) Toxic cardenolides: chemical ecology and coevolution of specialized plant-herbivore interactions. *New Phytologist* **194**, 28-45.
- Aldrich JR, Carroll SP, Lusby WR, *et al.* (1990) Sapindaceae, cyanolipids, and bugs. *Journal of Chemical Ecology* **16**, 199-210.
- Andersson MN, Binyameen M, Sadek MM, Schlyter F (2011) Attraction modulated by spacing of pheromone components and anti-attractants in a bark beetle and a moth. *Journal of Chemical Ecology* **37**, 899-911.
- Angelini DR, Kaufman TC (2004) Functional analyses in the hemipteran *Oncopeltus fasciatus* reveal conserved and derived aspects of appendage patterning in insects. *Developmental Biology* **271**, 306-321.
- Angelini DR, Kaufman TC (2005) Comparative developmental genetics and the evolution of arthropod body plans. *Annual Review of Genetics* **39**, 95-119.
- Angelini DR, Smith FW, Jockusch EL (2012) Extent With Modification: Leg Patterning in the Beetle *Tribolium castaneum* and the Evolution of Serial Homologs. *G3-Genes Genomes Genetics* **2**, 235-248.
- Arendt J, Reznick D (2008) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology & Evolution* **23**, 26-32.
- Arnold ML (1997) Natural Hybridization and Evolution. *Oxford University Press, Oxford*.
- Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge
- Bean AR (2007) A new system for determining which plant species are indigenous in Australia. *Australian Systematic Botany* **20**, 1-43.
- Becker S, Schneider H, Scheiner-Bobis G (2004) The highly conserved extracellular peptide, DSYG(893-896), is a critical structure for sodium pump function. *European Journal of Biochemistry* **271**, 3821-3831.
- Benton R, Sachse S, Michnick S, Vosshall L (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *Chemical Senses* **31**, A5-A5.
- Berres ME (2003) AFLP. Available from: <http://ravel.zoology.wisc.edu/sgaap/AFLP.html/AFLP.htm>.
- Bramer C, Dobler S, Deckert J, Stemmer M, Petschenka G (2015) Na⁺/K⁺-ATPase resistance and cardenolide sequestration: basal adaptations to host plant toxins in the milkweed bugs (Hemiptera: Lygaeidae: Lygaeinae). *Philosophical Transactions of the Royal Society B-Biological Sciences* **282**.
- Burgarella C, Lorenzo Z, Jabbour-Zahab R, *et al.* (2009) Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity* **102**, 442-452.
- Bush GL (1969) Sympatric Host Race Formation and Speciation in Frugivorous Flies of Genus *Rhagoletis* (Diptera, Tephritidae). *Evolution* **23**, 237-251.
- Butlin RK, Galindo J, Grahame JW (2008) Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B-Biological Sciences* **363**, 2997-3007.

- Caballero A, Quesada H (2010) Homoplasy and Distribution of AFLP Fragments: An Analysis In Silico of the Genome of Different Species. *Molecular Biology and Evolution* **27**, 1139-1151.
- Caballero A, Quesada H, Rolan-Alvarez E (2008) Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics* **179**, 539-554.
- Carroll SP (2007a) Brave New World: the epistatic foundations of natives adapting to invaders. *Genetica* **129**, 193-204.
- Carroll SP (2007b) Natives adapting to invasive species: Ecology, genes, and the sustainability of conservation. *Ecological Research* **22**, 892-901.
- Carroll SP (2008) Facing change: forms and foundations of contemporary adaptation to biotic invasions. *Molecular Ecology* **17**, 361-372.
- Carroll SP, Boyd C (1992) Host Race Radiation in the Soapberry Bug - Natural-History with the History. *Evolution* **46**, 1052-1069.
- Carroll SP, Dingle H, Famula TR (2003) Rapid appearance of epistasis during adaptive divergence following colonization. *Proceedings of the Royal Society B-Biological Sciences* **270**, S80-S83.
- Carroll SP, Dingle H, Famula TR, Fox CW (2001) Genetic architecture of adaptive differentiation in evolving host races of the soapberry bug, *Jadera haematoloma*. *Genetica* **112**, 257-272.
- Carroll SP, Dingle H, Klassen SP (1997) Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. *Evolution* **51**, 1182-1188.
- Carroll SP, Klassen STP, Dingle H (1998) Rapidly evolving adaptations to host ecology and nutrition in the soapberry bug. *Evolutionary Ecology* **12**, 955-968.
- Carroll SP, Loye J (2006) Invasion, colonization, and disturbance; historical ecology of the endangered Miami blue butterfly. *Journal of Insect Conservation* **10**, 13-27.
- Carroll SP, Loye JE (1987) Specialization of *Jadera* Species (Hemiptera, Rhopalidae) on the Seeds of Sapindaceae (Sapindales), and Coevolutionary Responses of Defense and Attack. *Annals of the Entomological Society of America* **80**, 373-378.
- Carroll SP, Loye JE (2012) Soapberry Bug (Hemiptera: Rhopalidae: Serinethinae) Native and Introduced Host Plants: Biogeographic Background of Anthropogenic Evolution. *Annals of the Entomological Society of America* **105**, 671-684.
- Carroll SP, Loye JE, Dingle H, *et al.* (2005a) And the beak shall inherit - evolution in response to invasion. *Ecology Letters* **8**, 944-951.
- Carroll SP, Loye JE, Dingle H, Mathieson M, Zalucki MP (2005b) Ecology of *Leptocoris* Hahn (Hemiptera : Rhopalidae) soapberry bugs in Australia. *Australian Journal of Entomology* **44**, 344-353.
- Carroll SP, Loye JE, Dingle H, Mathieson M, Zalucki MP (2006) Ecology of *Leptocoris* Hahn (Hemiptera : Rhopalidae) soapberry bugs in Australia (vol 44, pg 344, 2005). *Australian Journal of Entomology* **45**, 185-185.
- Chatfield MWH, Kozak KH, Fitzpatrick BM, Tucker PK (2010) Patterns of differential introgression in a salamander hybrid zone: inferences from

- genetic data and ecological niche modelling. *Molecular Ecology* **19**, 4265-4282.
- Chen S, Luetje CW (2012) Identification of new agonists and antagonists of the insect odorant receptor co-receptor subunit. *PLoS One* **7**, e36784.
- Cohen B, Simcox AA, Cohen SM (1993) Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. *Development* **117**, 597-608.
- Cooper CB, Arnot MI, Feng ZP, *et al.* (2000) Cross-talk between G-protein and protein kinase C modulation of N-type calcium channels is dependent on the G-protein beta subunit isoform. *Journal of Biological Chemistry* **275**, 40777-40781.
- Dalla S, Swarts HG, Koenderink JB, Dobler S (2013) Amino acid substitutions of Na,K-ATPase conferring decreased sensitivity to cardenolides in insects compared to mammals. *Insect Biochemistry and Molecular Biology* **43**, 1109-1115.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**, 772-772.
- de Jong MA, Wahlberg N, van Eijk M, Brakefield PM, Zwaan BJ (2011) Mitochondrial DNA signature for range-wide populations of *Bicyclus anynana* suggests a rapid expansion from recent refugia. *PLoS One* **6**, e21385.
- Deisig N, Giurfa M, Sandoz JC (2010) Antennal lobe processing increases separability of odor mixture representations in the honeybee. *Journal of Neurophysiology* **103**, 2185-2194.
- Despres L, David JP, Gallet C (2007) The evolutionary ecology of insect resistance to plant chemicals. *Trends in Ecology & Evolution* **22**, 298-307.
- Dingle H, Carroll SP, Famula TR (2009) Influence of genetic architecture on contemporary local evolution in the soapberry bug, *Jadera haematoloma*: artificial selection on beak length. *Journal of Evolutionary Biology* **22**, 2031-2040.
- Djamgoz MB, Ready PD, Billingsley PF, Emery AM (1998) Insect Na(+)/K(+)-ATPase. *Journal of Insect Physiology* **44**, 197-210.
- Dobler S, Dalla S, Wagschal V, Agrawal AA (2012) Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13040-13045.
- Ehrlich PR, Raven PH (1964) Butterflies and Plants - a Study in Coevolution. *Evolution* **18**, 586-608.
- Emelianov I, Simpson F, Narang P, Mallet J (2003) Host choice promotes reproductive isolation between host races of the larch budmoth *Zeiraphera diniana*. *Journal of Evolutionary Biology* **16**, 208-218.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.
- Faith DP, Cranston PS (1991) Could a Cladogram This Short Have Arisen by Chance Alone - on Permutation Tests for Cladistic Structure. *Cladistics-the International Journal of the Willi Hennig Society* **7**, 1-28.

- Feder JL, Berlocher SH, Roethele JB, *et al.* (2003a) Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10314-10319.
- Feder JL, Roethele FB, Filchak K, Niedbalski J, Romero-Severson J (2003b) Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* **163**, 939-953.
- Frenzel M, Brandl R (2003) Diversity and abundance patterns of phytophagous insect communities on alien and native host plants in the Brassicaceae. *Ecography* **26**, 723-730.
- Futuyma DJ, Agrawal AA (2009) Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18054-18061.
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* **21**, 394-407.
- Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology* **18**, 1207-1224.
- Gompert Z, Buerkle CA (2010) INTROGRESS: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources* **10**, 378-384.
- Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC (2006) Homoploid hybrid speciation in an extreme habitat. *Science* **314**, 1923-1925.
- Graves SD, Shapiro AM (2003) Exotics as host plants of the California butterfly fauna. *Biological Conservation* **110**, 413-433.
- Groman JD, Pellmyr O (2000) Rapid evolution and specialization following host colonization in a yucca moth. *Journal of Evolutionary Biology* **13**, 223-236.
- Gross GF (1960) A revision of the genus *Leptocoris* Hahn (Heteroptera: Coreidae: Rhopalinae) from the Indo-Pacific and Australian Regions. *Records of the Australian Museum* **13**, 403-451.
- Hallem EA, Dahanukar A, Carlson JR (2006) Insect odor and taste receptors. *Annual Review of Entomology* **51**, 113-135.
- Hansen O (1984) Interaction of Cardiac-Glycosides with (Na⁺ + K⁺)-Activated Atpase - a Biochemical Link to Digitalis-Induced Inotropy. *Pharmacological Reviews* **36**, 143-163.
- Hendry AP, Farrugia TJ, Kinnison MT (2008) Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* **17**, 20-29.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965-1978.
- Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* **83**, 189-195.
- Hoekstra HE, Hirschmann RJ, Bunday RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101-104.

- Horisberger JD (2004) Recent insights into the structure and mechanism of the sodium pump. *Physiology* **19**, 377-387.
- Hua JF, Zhang S, Cui JJ, *et al.* (2012) Identification and Binding Characterization of Three Odorant Binding Proteins and One Chemosensory Protein from *Apolygus lucorum* (Meyer-Dur). *Journal of Chemical Ecology* **38**, 1163-1170.
- Hubbs CL (1940) Speciation of fishes. *American Naturalist* **74**, 198-211.
- Ishimaru Y, Nakamura T, Bando T, *et al.* (2015) Involvement of dachshund and Distal-less in distal pattern formation of the cricket leg during regeneration. *Scientific Reports* **5**, 8387.
- Janz N, Nylin S, Wahlberg N (2006) Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology* **6**.
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**.
- Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* **101**, 92-103.
- Jones WD, Nguyen TAT, Kloss B, Lee KJ, Vosshall LB (2005) Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Current Biology* **15**, R119-R121.
- Karban R, Agrawal AA (2002) Herbivore offense. *Annual Review of Ecology and Systematics* **33**, 641-664.
- Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780.
- Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* **11**, 852-866.
- Koopman WJM (2005) Phylogenetic signal in AFLP data sets. *Systematic Biology* **54**, 197-217.
- Kosman E, Leonard KJ (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology* **14**, 415-424.
- Kuusela J, Zietara MS, Lumme J (2007) Hybrid origin of Baltic salmon-specific parasite *Gyrodactylus salaris*: a model for speciation by host switch for hemiclinal organisms. *Molecular Ecology* **16**, 5234-5245.
- Lambrinos JG (2004) How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology* **85**, 2061-2070.
- Langerhans RB, Gifford ME, Joseph EO (2007) Ecological speciation in *Gambusia* fishes. *Evolution* **61**, 2056-2074.
- Lavergne S, Mouquet N, Thuiller W, Ronce O (2010) Biodiversity and Climate Change: Integrating Evolutionary and Ecological Responses of Species and Communities. *Annual Review of Ecology, Evolution, and Systematics* **41**, 321-350.

- Leal WS (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annual Review of Entomology* **58**, 373-391.
- Ley AC, Hardy OJ (2014) Contrasting patterns of gene flow between sister plant species in the understory of African moist forests - The case of sympatric and parapatric Marantaceae species. *Molecular Phylogenetics and Evolution* **77**, 264-274.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.
- Lynch M, Milligan BG (1994) Analysis of Population Genetic-Structure with Rapid Markers. *Molecular Ecology* **3**, 91-99.
- Mardon G, Solomon NM, Rubin GM (1994) dachshund encodes a nuclear protein required for normal eye and leg development in Drosophila. *Development* **120**, 3473-3486.
- Matsubayashi KW, Ohshima I, Nosil P (2010) Ecological speciation in phytophagous insects. *Entomologia Experimentalis Et Applicata* **134**, 1-27.
- McBride CS (2007) Rapid evolution of smell and taste receptor genes during host specialization in Drosophila sechellia. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4996-5001.
- Mercader RJ, Aardema ML, Scriber JM (2009) Hybridization leads to host-use divergence in a polyphagous butterfly sibling species pair. *Oecologia* **158**, 651-662.
- Mitter C, Farrell B, Wiegmann B (1988) The Phylogenetic Study of Adaptive Zones - Has Phytophagy Promoted Insect Diversification. *American Naturalist* **132**, 107-128.
- Navarro A, Barton NH (2002) The effects of multilocus balancing selection on neutral variability. *Genetics* **161**, 849-863.
- Nei M, Li WH (1979) Mathematical-Model for Studying Genetic-Variation in Terms of Restriction Endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 5269-5273.
- Novotny V, Miller SE, Leps J, *et al.* (2004) No tree an island: the plant-caterpillar food web of a secondary rain forest in New Guinea. *Ecology Letters* **7**, 1090-1100.
- Nyman T, Vikberg V, Smith DR, Boeve JL (2010) How common is ecological speciation in plant-feeding insects? A 'Higher' Nematinae perspective. *BMC Evolutionary Biology* **10**.
- Palczewski K, Kumasaka T, Hori T, *et al.* (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745.
- Palumbi SR (2001) Evolution - Humans as the world's greatest evolutionary force. *Science* **293**, 1786-1790.
- Panganiban G, Irvine SM, Lowe C, *et al.* (1997) The origin and evolution of animal appendages. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 5162-5166.
- Paris M, Bonnes B, Ficetola GF, Poncet BN, Despres L (2010) Amplified fragment length homoplasy: in silico analysis for model and non-model species. *BMC Genomics* **11**.
- Patil AG, Joshi KA, Patil DA, Phatak AV, Chandra N (2010) Pharmacognostic and physico-chemical studies on the leaves of *Cardiospermum halicacabum* L. *Pharmacognosy Journal*, p267.

- Peccoud J, Ollivier A, Plantegenest M, Simon JC (2009) A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7495-7500.
- Petschenka G, Fandrich S, Sander N, *et al.* (2013) Stepwise Evolution of Resistance to Toxic Cardenolides Via Genetic Substitutions in the Na(+)/K+-Atpase of Milkweed Butterflies (Lepidoptera: Danaini). *Evolution* **67**, 2753-2761.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231-259.
- Prpic NM, Tautz D (2003) The expression of the proximodistal axis patterning genes *Distal-less* and *dachshund* in the appendages of *Glomeris marginata* (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages. *Developmental Biology* **260**, 97-112.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**, 2092-2100.
- Raza SA, Hussain S, Riaz H, Mahmood S (2013) Review of beneficial and remedial aspect of *Cardiospermum halicacabum*. *African Journal of Pharmacy and Pharmacology* **7**, 3026-3033.
- Reeves PA, Richards CM (2009) Accurate Inference of Subtle Population Structure (and Other Genetic Discontinuities) Using Principal Coordinates. *PLoS One* **4**.
- Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**, 359-389.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.
- Rosenblum EB, Hoekstra HE, Nachman MW (2004) Adaptive reptile color variation and the evolution of the *Mc1r* gene. *Evolution* **58**, 1794-1808.
- Rospars JP, Gremiaux A, Jarriault D, *et al.* (2014) Heterogeneity and convergence of olfactory first-order neurons account for the high speed and sensitivity of second-order neurons. *PLoS Computational Biology* **10**, e1003975.
- Roux C, Pauwels M, Ruggiero MV, *et al.* (2013) Recent and ancient signature of balancing selection around the S-locus in *Arabidopsis halleri* and *A. lyrata*. *Molecular Biology and Evolution* **30**, 435-447.
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters* **8**, 336-352.
- Sato K, Pellegrino M, Nakagawa T, *et al.* (2008) Insect Olfactory Receptors are Heteromeric Ligand-Gated Ion Channels. *Chemical Senses* **33**, S74-S74.
- Schluter D (2009) Evidence for Ecological Speciation and Its Alternative. *Science* **323**, 737-741.
- Schluter D, Nagel LM (1995) Parallel Speciation by Natural-Selection. *American Naturalist* **146**, 292-301.
- Schwarz D, Matta BM, Shakir-Botteri NL, McPheron BA (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* **436**, 546-549.
- 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. *Annual Review of Ecology Evolution and Systematics* **37**, 545-579.

- Simon KS, Townsend CR (2003) Impacts of freshwater invaders at different levels of ecological organisation, with emphasis on salmonids and ecosystem consequences. *Freshwater Biology* **48**, 982-994.
- Strauss SY, Lau JA, Carroll SP (2006) Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecology Letters* **9**, 354-371.
- Suresh S, Rathishkumar S, Rajeshwari V, *et al.* (2012) Phytochemical analysis and antibacterial potential of *Cardiospermum halicacabum* Linn. (Sapindaceae) *International Journal Of Pharmacy & Life Sciences* **3**, 2209-2212.
- Swofford DL (1999) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). *Sinauer Associates, Sunderland, Massachusetts*.
- Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony(* and Other Methods). *Sinauer Associates, Sunderland, Massachusetts*.
- Tajima F (1989) Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **123**, 585-595.
- Taylor RW, Romaine IM, Liu C, *et al.* (2012) Structure-Activity Relationship of a Broad-Spectrum Insect Odorant Receptor Agonist. *ACS Chemical Biology* **7**, 1647-1652.
- Thomas Y, Bethenod MT, Pelozuelo L, Frerot B, Bourguet D (2003) Genetic isolation between two sympatric host-plant races of the European corn borer, *Ostrinia nubilalis* Hubner. I. sex pheromone, moth emergence timing, and parasitism. *Evolution* **57**, 261-273.
- Toju H (2009) Natural selection drives the fine-scale divergence of a coevolutionary arms race involving a long-mouthed weevil and its obligate host plant. *BMC Evolutionary Biology* **9**.
- Vaha JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63-72.
- Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasmy and of a relationship between degree of homoplasmy and fragment size. *Molecular Ecology* **11**, 139-151.
- Via S (1991a) The Genetic-Structure of Host Plant Adaptation in a Spatial Patchwork - Demographic Variability among Reciprocally Transplanted Pea Aphid Clones. *Evolution* **45**, 827-852.
- Via S (1991b) Specialized Host Plant Performance of Pea Aphid Clones Is Not Altered by Experience. *Ecology* **72**, 1420-1427.
- Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9939-9946.
- Vos P, Hogers R, Bleeker M, *et al.* (1995) Aflp - a New Technique for DNA-Fingerprinting. *Nucleic Acids Research* **23**, 4407-4414.
- Winkler KA, Pamminger-Lahnsteiner B, Wanzenbock J, Weiss S (2011) Hybridization and restricted gene flow between native and introduced stocks of Alpine whitefish (*Coregonus* sp.) across multiple environments. *Molecular Ecology* **20**, 456-472.
- Wood TK, Keese MC (1990) Host-Plant-Induced Assortative Mating in *Enchenopa* Treehoppers. *Evolution* **44**, 619-628.

- Yang X, Weber M, Zarinkamar N, *et al.* (2009) Probing the *Drosophila* retinal determination gene network in *Tribolium* (II): The Pax6 genes *eyeless* and *twin of eyeless*. *Developmental Biology* **333**, 215-227.
- Yoda A, Yoda S (1982) Interaction between Ouabain and the Phosphorylated Intermediate of Na,K-ATPase. *Molecular Pharmacology* **22**, 700-705.
- Yu Y, Andres JA (2014) Genetic Architecture of Contemporary Adaptation to Biotic Invasions: Quantitative Trait Locus Mapping of Beak Reduction in Soapberry Bugs. *G3-Genes Genomes Genetics* **4**, 255-264.
- Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P (2012) Parallel Molecular Evolution in an Herbivore Community. *Science* **337**, 1634-1637.
- Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* **8**, 907-913.