2014

Character Evolution and Microbial Community Structure in a Host-associated Grasshopper

Tyler Raszick
University of Central Florida

CHARACTER EVOLUTION AND MICROBIAL COMMUNITY STRUCTURE IN A HOST-ASSOCIATED GRASSHOPPER

by

TYLER JAY RASZICK
B.S. University of Florida, 2010

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2014

ABSTRACT

The spotted bird grasshopper, *Schistocerca lineata* Scudder (Orthoptera: Acrididae), is a widely distributed species found throughout most of the continental United States and southern Canada. This species is known to be highly variable in morphology, with many distinct ecotypes across its native range. These ecotypes display high levels of association with type-specific host plants. Understanding the evolutionary relationships among different ecotypes is crucial groundwork for studying the process of ecological differentiation. I examine four ecotypes from morphological and phylogeographic perspectives, and look for evidence of distinct evolutionary lineages within the species. I also begin to explore the potential role of the microbial community of these grasshoppers in ecological divergence by using 454 pyrosequencing to see if the microbial community structure reflects the ecology of the grasshoppers. I find support for a distinct aposematic lineage when approaching the data from a phylogeographic perspective and also find that this ecotype tends to harbor a unique bacterial community, different from that of a single other ecotype.

ACKNOWLEDGMENTS

I would like to acknowledge my advisor H. Song of the University of Central Florida for mentorship throughout my degree, as well as my thesis committee, K. Fedorka and E. Hoffman. G. Sword and P. Lenhart of Texas A&M University was invaluable, providing guidance throughout the project, and assisting with collection efforts. I am grateful to C. R. Bomar of the University of Wisconsin – Stout for providing WI specimens. I thank S. Gotham for assisting with 2011 collection efforts and G. Alava for carrying out PCR amplifications. C. Gale was a trooper, and continued to help throughout the 2013 collecting effort despite suffering parasitism by chiggers.

This thesis was funded by NSF Grant DEB-1064082 and IOS-1253493 to Hojun Song and by the Orthopterists' Society Student Research Fund. Travel support provided by the University of Central Florida Student Government Association.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I: GENERAL INTRODUCTION

Evolutionary Biology and Ecological Differentiation

To achieve mechanistic knowledge of the interactions between ecology and evolutionary processes is a core aim of evolutionary biology, and is fundamental to understanding the process of speciation at the microevolutionary scale [1,2]. Within a single species, populations may diverge along a continuum on which the strength of reproductive barriers to gene flow can vary from weak to strong [2,3,4,5]. These barriers may arise from postzygotic regulation such as hybrid inviability or hybrid breakdown, or from prezygotic sources such as temporal or geographic isolation. Although a variety of prezygotic mechanisms have been proposed as potential routes to reproductive isolation, many of these processes remain poorly understood in complex systems [3,6,7,8].

One such mechanism is ecological differentiation. Ecological differentiation is a divergent process in which disruptive selection from environmental differences or ecological interactions act in contrasting directions on two or more populations of a single species, and can potentially lead to reproductive isolation. Fundamental insights about ecological differentiation, especially in insects, have come from studies of herbivorous species with ecologically divergent host races, sometimes referred to as ecotypes [4,5,9,10]. Host-associated differentiation (HAD) is a specific case of ecological differentiation in which host-associated taxa develop divergent traits across populations due to associations with different hosts. This may be especially common in herbivorous insects, and numerous studies have investigated this phenomenon in a variety of taxa. Some *Drosophila* spp. (Diptera: Drosophilidae), as well as numerous lepidopterans, have larval stages that are dependent upon a single host plant species, while others

have a complex of hosts that they might utilize [11]. Beetles of the genus *Neochlamisus* (Coleoptera: Chrysomelidae) exhibit species-specific host associations, with each species feeding, growing, mating, and laying eggs on a single host plant throughout its lifetime [11]. *Neochlamisus bebbianae*, however, seems to be at a different point along a divergence continuum when compared to the rest of the genus. Different populations display distinct host-forms, each associated with one of six plant genera, across five families [12]. Each host-form can be ecologically and genetically differentiated from the others [12]. In *Timema cristinae* (Phasmatodea: Timemidae), host plant association has led to the development of two ecotypes, each with distinct coloration and morphology, adapted for crypsis on a certain host plant genus [13]. An individual of one ecotype experiences increased predation pressure by being active on the "wrong" host plant [14]. This may represent a case of incomplete speciation, and it has been further suggested that this could be a stage of an ongoing speciation event [15,16].

System of Study: Schistocerca lineata

The spotted bird grasshopper *Schistocerca lineata* Scudder, 1899 (Orthoptera: Acrididae) is a widely distributed North American species that occurs in highly localized and potentially isolated populations that are often associated with different host plants [17,18,19,20]. This species presents a unique opportunity to study patterns of ecological differentiation at the microevolutionary scale because of its unique ecology and population-level diversity. Some populations have been shown to display ontogenetic specialization in which juveniles are host specific, but adults become more generalist [19,20], and this is likely the general pattern for the species. A recent revision of the Alutacea Group within the genus *Schistocerca* described four ecotypes of *S. lineata*; (1) aposematic, (2) brown, (3) typical, (4) olive-green; mostly based on

external coloration of museum-preserved adult specimens [17]. Each ecotype is associated with a particular phenotype, the most obvious component of which is color (Fig. 1).

Figure 1: Ecotypes of *S. lineata*
Clockwise from top-left: (1) aposematic; (2) brown; (3) typical; (4) olive-green.
Photo credits: (1) DiBurro, 2013; (2) Haarstad, Baumert, Sheps, Treon, 2002; (3) Cotinis, 2011; (4) Plagens, 2008.

Out of these four ecotypes, only two, both occurring in Texas, have well-characterized ecological traits and life histories. Nymphs of the aposematic ecotype feed on the toxic wafer ash, *Ptelea trifoliata* (L.) and display a characteristic yellow and black pattern (Fig. 1). These *Ptelea*-feeders exhibit density-dependent aposematism and derive chemical defense from the host plant [21,22]. The typical ecotype is cryptic in color and is known to be associated with dewberry, *Rubus*

trivialis Michx. in Texas [17,18,19,20]. For these two ecotypes, nymphal survival is strongly correlated with the availability of the preferred host plant species, and the two groups have been shown to be genetically distinct based on two mtDNA markers [19]. The olive-green ecotype is found in Colorado and known to be associated with invasive saltcedar, *Tamarix ramosissima* Ledeb., which was introduced to the region in the early 1900s [18,23]. The brief history of this plant in the United States suggests that Colorado populations of *S. lineata* must have developed this association extremely rapidly, but no empirical study has evaluated the strength of the association nor tested the hypothesis that this is a genetically distinct group. Not much is known about the ecology of the brown ecotype.

Other investigations into the ecology of *S. lineata* indicate that there may be a wide variety of distinct host plants with which at least one population of *S. lineata* is associated suggesting that, for some populations, the ecotype designation may be of limited value [17,24]. The overwhelming intraspecific morphological variation found in *S. lineata* demands further investigation, and provides an excellent opportunity to simultaneously study patterns of evolution in a rapidly evolving host-associated, hemimetabolous insect. In Chapter II, I characterize genetic differentiation and identify patterns of morphological divergence across the four ecotypes.

Microbial Community Structure

In Chapter III, I begin to investigate the potential for a role of symbiotic bacteria in the divergence of ecotypes. Symbiotic bacteria can have a significant influence on the fitness of host organisms, and may play an important role in speciation events involving their hosts [25,26,27]. Nonetheless, investigation of the role of symbiotic bacteria in the processes of speciation has not

yet found its way into the mainstream of speciation research, though there are examples [25,26,27]. Here, I compare the microbial community structure to the known ecology of the two most well understood ecotypes of *S. lineata*. It has been shown in the congeneric *S. gregaria* that the gut microbial community is derived from the diet, so it stands to reason that two ecotypes with divergent host plant associations might harbor different bacterial communities. Using next-generation sequencing (NGS) to quantify the abundances of bacterial genera in eight populations, I illuminate a pattern of microbial divergence that reflects both ecological and genetic divergence.

CHAPTER II: CHARACTER EVOLUTION IN A HOST-ASSOCIATED GRASSHOPPER

Introduction

Many insect species exist as numerous divergent populations [4,5,10,19, 28]. Sometimes, these populations are ecologically and morphologically distinct from each other, and the term "ecotype" is often used to denote each distinct population. The concept of ecotype dates back to Turesson (1922) who proposed that a species could exist as many ecological units that arise as a result of the genetic response to a particular habitat. Since then, the term ecotype has been loosely used throughout the literature. While conceptually straightforward, the delimitation of ecotypes is actually often difficult and can be quite subjective. For example, if a population of an herbivorous insect species has evolved a preference for a particular host plant, and if it is possible to readily recognize this ecological characteristic, one may use the term ecotype for this population without much hesitation. Similarly, if another population of the same species always displays a certain phenotype and seems to be found in a particular habitat, one might also use the term ecotype to denote this population without any regards to feeding biology. Thus, the ecotype designation does not necessarily imply a certain evolutionary process (selection, drift, or demography), but it does suggest that a particular ecotype is on its own evolutionary trajectory, distinct from other ecotypes. The divergence of ecotypes fits well within the concept of speciation continuum [4,5]. If the ecotype continues to persist, it may potentially become an incipient species in the context of ecological speciation [28].

Ecotypes are recognizable by ecological, phenotypic, or genetic traits, or by any combination of these three traits. Often, multiple ecotypes exist for a given species, but there does not seem to be an established explanation that describes a general pattern for the divergence

of those traits that are used to recognize these ecotypes. Although several studies have examined host-race or ecotype formation in insects, very few have attempted to describe the patterns of character divergence during the very early stages of ecotype divergence [4,5,10,29]. At this stage, it is expected that a so-called "mosaic" pattern of character divergence should emerge. Mosaic evolution describes a phenomenon when different characters show a variety of different patterns of evolution, which is often invoked in paleontological studies [30,31,32]. Although the concept originated from a macroevolutionary perspective, it is possible to apply at a microevolutionary scale [18,33]. For instance, mosaic patterns can be found when characters diverge differentially across multiple populations or ecotypes. Incomplete lineage sorting is a result of mosaic evolution of molecular characters in the sense that individual gene loci evolve at different rates resulting in incongruence among those loci. Similar patterns have been observed in morphological data from a variety of taxa [18,34,35]. However, there has not been much attention paid to understanding the mosaic patterns of character divergence during the early stages of ecological differentiation.

In this study, I describe patterns of character divergence in several known ecotypes of a rapidly-diverging polyphagous grasshopper species. Using a variety of morphological and molecular markers, I assess phenotypic and genetic divergence in four ecotypes of *Schistocerca lineata*. I sample both size-dependent and size-independent morphological characters to evaluate how morphology diverges across ecotypes. Since different host plant species might confer differing nutrition to the insect, I hypothesize that size will vary across ecotypes. Likewise, there is known population-level divergence in male sexual structures [18]. Thus, I hypothesize that there will be ecotype-level variation in male cerci, which have been shown to be appropriate for assessing population-level variation in grasshoppers [18]. Considering that different selection

pressures likely exist with regards to ecology and sexual selection, we can expect that these two different types of characters may diverge at different rates, exhibiting a mosaic pattern. Alternatively, if selection is weak, a mosaic pattern could arise due simply to drift. It could also happen that no difference is seen in either set of characters. If this were the case, then we would expect to see genetic evidence of population panmixis. Two ecotypes of *S. lineata* have previously been shown to be genetically distinct based on partial sequences of the mitochondrial 16S and 12S rRNA [19]. I expand both taxon and character sampling to all four described ecotypes of *S. lineata,* and relate the pattern of genetic divergence back to the patterns morphological and ecological divergence across ecotypes.

Methods

Taxon Sampling

Specimens of *S. lineata* were collected from eight localities across the continental United States (Table 1). Specifically, I sampled two populations from Texas and one each from Oklahoma and Colorado. Museum and alcohol preserved specimens from another Oklahoma population and from another Colorado population, as well as specimens from a Kansas locality were contributed by H. Song. Adults and nymphs of varied ages were collected using hand capture and sweep netting techniques. Animals were kept alive, transported to the Song Laboratory of Insect Systematics and Evolution at UCF, and reared to adulthood in moderate-density colonies on a 16:8 hour day-night cycle in a temperature-controlled rearing room at ~26.7ºC. Grasshoppers were fed a diet of Romaine lettuce and wheat bran, supplemented with known or assumed host plants collected from the field *ad libitum*. The eighth collecting locality (WI) was sampled by a colleague (C. Bomar) who provided the adult specimens preserved 100%

ethanol. Colony-reared specimens were also preserved in 100% ethanol upon natural death and stored at –80ºC. A total of 109 individuals were sampled for use in morphological and molecular analyses.

Morphological Character Sampling

In this study, I quantified two types of morphological traits (size-dependent linear measurements and size-independent shapes) across the eight populations of *S. lineata* to measure population-level divergence.

Size-dependent Morphology: Hind femur length (F), head width (C), and pronotum length (P), have been shown to be good metrics for evaluating overall individual size in grasshoppers [18]. For hind femur measurements, the entire left leg of the grasshopper was removed and laid flat on a standardized ceramic plate. If the left leg was not present, or had been potentially altered by DNA extraction, the right leg was used instead. This does not present an issue because insects are bilaterally symmetrical, and both legs should be identical in length. Deformed or poorly developed grasshoppers were not used in this analysis. The legs were photographed using the high resolution BK Plus Lab Imaging System (Visionary Digital). Femur measurements were taken in Adobe Photoshop CS5, calibrated for the appropriate lens (50mm, f-stop 8.0, zoom 1:6) and camera (Canon EOS 7D) settings. For the purposes of this study, femur length was considered to be from the apex of the curve at the base of the femur to the notch where the femur meets the tibia at the joint. Measurements were taken three times and averaged for a final value in order to minimize human error. To obtain values for C and P, I used Mitutoyo ABSOLUTE Digimatic digital calipers to manually measure the head at its widest point, below the eyes, and to measure the length of the pronotum. Again, measurements were taken three times, and the values averaged.

Table 1: Collection summary and descriptions of adult phenotypes

Population (abbr.)	Date	Host Plant	Habitat	Adult Coloration (ecotype)
Altair, TX (TX-RT)	July 2011	*Rubus trivialis*	dense shrubs and grasses	yellow-brown with yellow dorsal stripe, some with green markings, black hind tibiae (typical)
Austin, TX (TX-PT)	July 2011	*Ptelea trifoliata*	woody area, some canopy	bright yellow with yellow dorsal stripe and blue eyes; black markings and hind tibiae (aposematic)
Fort Sill, OK (OK-FS)	July 2004	unknown	short grass, mixed	sandy yellow with black lateral stripes and yellow dorsal stripe and black hind tibiae (typical)
Fort Cobb, OK (OK-FC)	July 2011	unknown	dry, sandy	rusty to dark brown with yellow dorsal stripe and brown or black hind tibiae (typical)
John Martin Reservoir, CO (CO-JMR)	July 2004	*Tamarix ramosissima*	dry, sandy	olive green with yellow dorsal stripe and red hind tibiae (olive-green)
Blue Lake, CO (CO-BL)	July 2011	*Tamarix ramosissima*	grassy lakeshore	pale green with yellow dorsal stripe and red hind tibiae (olive-green)
Hollister Wildlife Area, KS (KS)	July 2004	*Rhus* sp.	tall grass prairie	dark brown with yellow dorsal stripe and black hind tibiae (typical)
Wisconsin (WI)	August 2005	unknown	N/A	brown, some lacking yellow dorsal stripe and brown or black hind tibae (brown)

Values of F, C, and P are highly correlated metrics of grasshopper size. Although any one of these measures may have been suitable for direct comparisons across populations, I analyzed all three metrics simultaneously via principal components analysis (PCA) using JMP 10.0.0 (SAS Institute Inc., 2012). I used equal numbers of males and females whenever possible. Mean PC 1 scores were compared across populations using ANOVA (Table 2), and Tukey's HSD test was used to identify significantly differences between ecotypes (Fig. 2).

Size-independent Morphology: The shapes of male cerci may provide an indicator of population-level divergence that is independent of individual grasshopper size [18]. To prepare for images for a morphometric analysis, the left cercus of each male was dissected from the specimen and photographed on a standardized petri dish-and-slide mount (n = 65). High-resolution photographs of the cerci were taken using a 100mm lens, with an f-stop of 5.6 and zoom 1:1.5. I used the software in the TPS series to carry out a landmark-based geometric morphometric analysis [36]. I created the input files using tpsUtil and quantified the shape of male cercus as a set of landmark coordinates (5 Type-II landmarks and 21 semi-sliding Type-III landmarks) using tpsDig2 [36]. I calculated partial warp scores for each specimen based on the landmark data and performed a relative warp analysis (a type of PCA) based on the resulting partial warp scores using tpsRelW [36]. Deviations of individuals' structures from the consensus shape are expressed as relative warp (RW) scores (Table 3). These RW scores were then compared for ecotype-level differences using MANOVA and Tukey's HSD test in JMP (Table 4). R 2.15.1 (The R Foundation for Statistical Computing, 2012) was used plot RW scores as functions of each other, and Inkscape 0.48.4 (The Inkscape Team) was used to draw minimum convex polygons (MCPs) representing the occupied morphospace for cerci in each ecotype. To visualize the shape variation in the morphospace along the resulting RW scores, I used tpsDig2 [36].

Molecular Character Sampling

I generated mitochondrial DNA (mtDNA) sequences from 70 specimens (including 7 outgroups; 2 *S. damnifica*, 5 *S. shoshone*) to use for phylogeographic analysis. Raw genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen) and sequences were amplified using standard PCR. I developed novel specific primers to amplify *cytochrome oxidase* subunit I (COI) in addition to the 12S and 16S genes utilized by Dopman *et al*. (2002), which have all been used to recover relatively recent signals of divergence across a variety of taxa. The transfer RNA (tRNA) valine, which occurs between 12S and 16S, was also amplified incidentally during PCR, and was also included in my analysis. PCR product was cleaned of residual primers and reagents using PrepEase 96-well plates (USB Corporation). Purified PCR product was sequenced at the BYU DNA Sequencing Center and proofread for quality using Sequencher 4.10.1 (Gene Codes Corporation, 2010). The resulting sequences were deposited at GenBank with accession numbers XXX-XXX.

Alignment and Phylogeographic Analysis: COI was aligned based on the conservation of amino acid sequence using MUSCLE in MEGA 5 [37]. Because rRNA and tRNA contain distinct, structurally functional stem and loop regions, I considered the secondary folding structure of these genes before building their alignments [38]. I used RNAfold (Institute for Theoretical Chemistry) to establish consensus structures for each gene, and RNAsalsa was used to align the sequences [39]. rRNA and tRNA alignments were then each divided into stem and loop subsets because base pairs in these structures may evolve at different rates, and COI was divided into three subsets based on codon position [38]. My final dataset had a total of nine subsets and a Bell number of 21,147, which represents the number of possible data partitioning schemes. PartitionFinder was used to compare every possible scheme with a Bayesian Information

Criterion (BIC) score, based on likelihood scores [38]. Besides recommending the optimal partitioning scheme, PartitionFinder also recommends the best-fit model of nucleotide evolution for each subset of the partition. Based on the partitioning scheme and the model of evolution recommended by the software, I reconstructed a phylogeny of *S. lineata* in a Bayesian framework using the hybrid MrBayes 3.1.2 build via the CIPRES Science Gateway [40,41]. I used default priors for the Bayesian analysis. Four runs with four chains each were carried out for 40 million generations, sampling every 2500 generations. Convergence was assessed with Tracer, and the first 25% of each run was discarded as burn-in. I also generated individual gene trees and trees for pairs of genes to assess the robustness of my topology using the same tree-building methods outlined above.

The same dataset was also reassessed using an alternative methodology to improve the robustness of my analyses. My concatenated dataset was trimmed to exclude uninformative sites or loci with too much missing information using the Phylogeny.fr Gblocks 0.91b utility, allowing for gaps within the final blocks [42,43]. The output was then analyzed using JModelTest to determine appropriate models of nucleotide evolution for each of three partitions (COI, 16S, and 12S; Valine was excluded from this analysis) [44,45]. A second series of Bayesian haplotype trees were generated from this trimmed dataset using default priors using MrBayes via CIPRES.

I generated a partial 16S gene tree using data from Dopman *et al.* (2002) mined from GenBank (accession numbers AF155548–AF155566) in addition to my own sequence data. My sequences were trimmed to reflect that dataset, and then the datasets were combined for use in a Bayesian phylogenetic analysis informed by PartitionFinder.

Results

Size-dependent Morphology

PC 1 explained 96.4% of the observed variation in the data, and is interpreted as an effective value of "grasshopper size." My model identified significant differences in PC 1 across the dataset, with significant predictor variables including gender and population nested within ecotype, as well as a significant effect of ecotype (Table 2).

Table 2: ANOVA summary for PC 1

Response = Grasshopper Size (PC 1)			
	d.f.	F Ratio	p value
MODEL	8	102.3567	< **0.0001**
Gender	1	544.8575	< **0.0001**
Pop[Ecotype]	4	36.3729	< **0.0001**
Ecotype	3	38.1115	< **0.0001**

Significant differences in grasshopper size across ecotypes were identified by the ANOVA, and the Tukey's HSD test revealed that grasshoppers of the brown ecotype were significantly small than grasshoppers of other ecotypes (Fig. 2). The olive-green ecotype was found to be of intermediate size, being significantly larger than the brown ecotype, but smaller than either of the other two ecotypes (Fig. 2). The aposematic and typical ecotypes were not found to be significantly different from each other with regards to size, but both ecotypes were significantly larger than both the brown and olive-green ecotypes (Fig. 2).

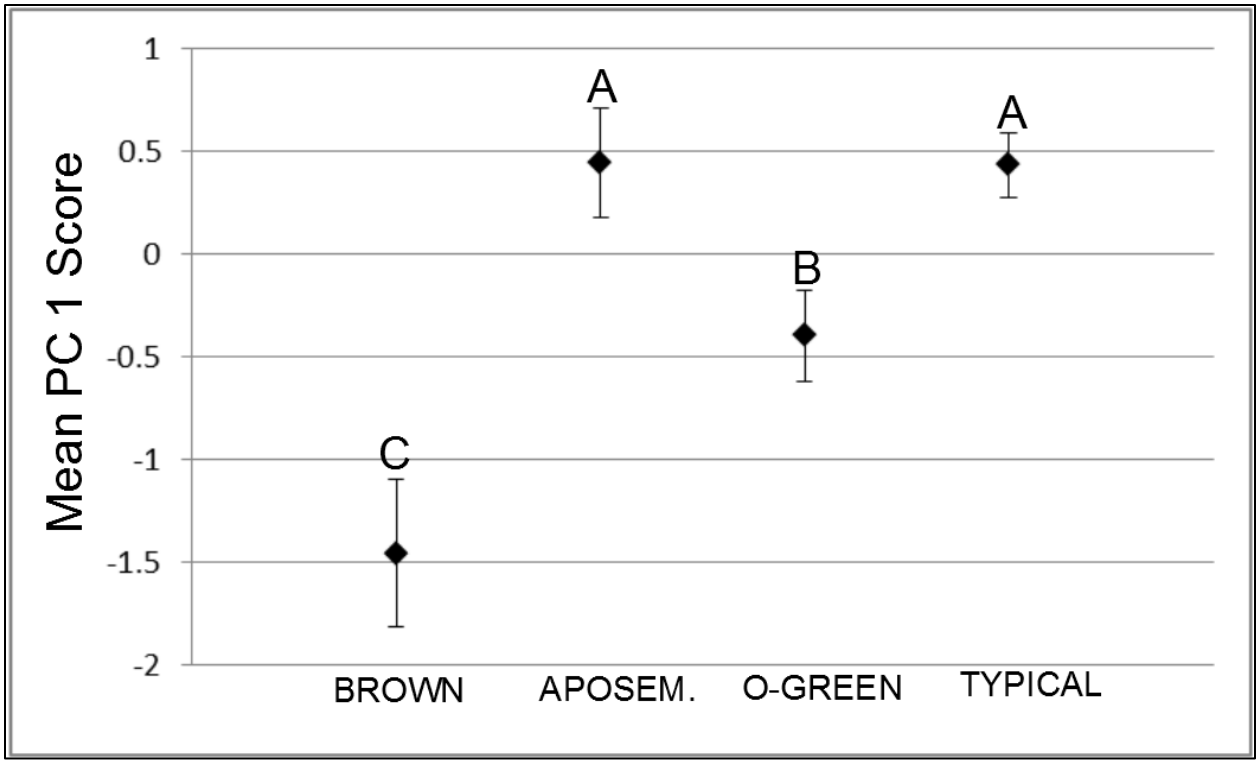

Figure 2: Mean size (PC 1) by ecotype
Means are least squared means. Error bars represent 95% confidence intervals and capital letters are Tukey's groupings for ecotype.

Size-independent Morphology

Relative warp analysis summarized the variation in male cerci shapes as 48 RW scores. The first six RW scores explained more than 90% of the overall variation (Table 3).

Table 3: Summary of RW analysis

RW	SV	% Var.	Cum. %
1	0.33396	37.68%	37.68%
2	0.23320	18.37%	56.05%
3	0.21146	15.11%	71.16%
4	0.17552	10.41%	81.57%
5	0.12601	5.36%	86.93%
6	0.09818	3.26%	90.19%
7	0.08246	2.30%	92.49%
8	0.07430	1.87%	94.35%
9	0.06023	1.23%	95.58%
10	0.05200	0.91%	96.49%

SV = Singular Value. Second column shows the percent of variation explained by each relative warp (RW), and the third column shows the cumulative percent variation explained.

My MANOVA model identified significant differences across the dataset for the first three RW scores, which cumulatively explain 71.16% of the observed variation. Although the variation in RW 1 was explained by the effect of population nested within ecotype, variation in RW 2 and RW 3 was due to the effect of ecotype (Table 4).

Table 4: MANOVA summary for first 6 RW scores

Response = RW 1				Response = RW 2			
	d.f.	F Ratio	p value		d.f.	F Ratio	p value
MODEL	7	2.2256	**0.0452**	MODEL	7	5.4539	< **0.0001**
Pop[Ecotype]	4	3.5181	**0.0123**	Pop[Ecotype]	4	0.6792	0.6092
Ecotype	3	0.3747	0.7716	Ecotype	3	11.4574	< **0.0001**
Response = RW 3				Response = RW 4			
	d.f.	F Ratio	p value		d.f.	F Ratio	p value
MODEL	7	2.3912	**0.0324**	MODEL	7	1.1128	0.3678
Pop[Ecotype]	4	1.875	0.1273	Pop[Ecotype]	4	1.3745	0.2542
Ecotype	3	3.2134	**0.0295**	Ecotype	3	0.7706	0.5152
Response = RW 5				Response = RW 6			
	d.f.	F Ratio	p value		d.f.	F Ratio	p value
MODEL	7	1.0614	0.4	MODEL	7	1.0073	0.4359
Pop[Ecotype]	4	0.2913	0.8825	Pop[Ecotype]	4	1.2351	0.3063
Ecotype	3	2.0694	0.1144	Ecotype	3	0.8483	0.4733

A Tukey's HSD test for RW 2 indicated that the olive-green ecotype was significantly different in shape from the other ecotypes, while the Tukey's test for RW 3 identified the aposematic ecotype as the significantly distinct group (Figs. 3 & 4).

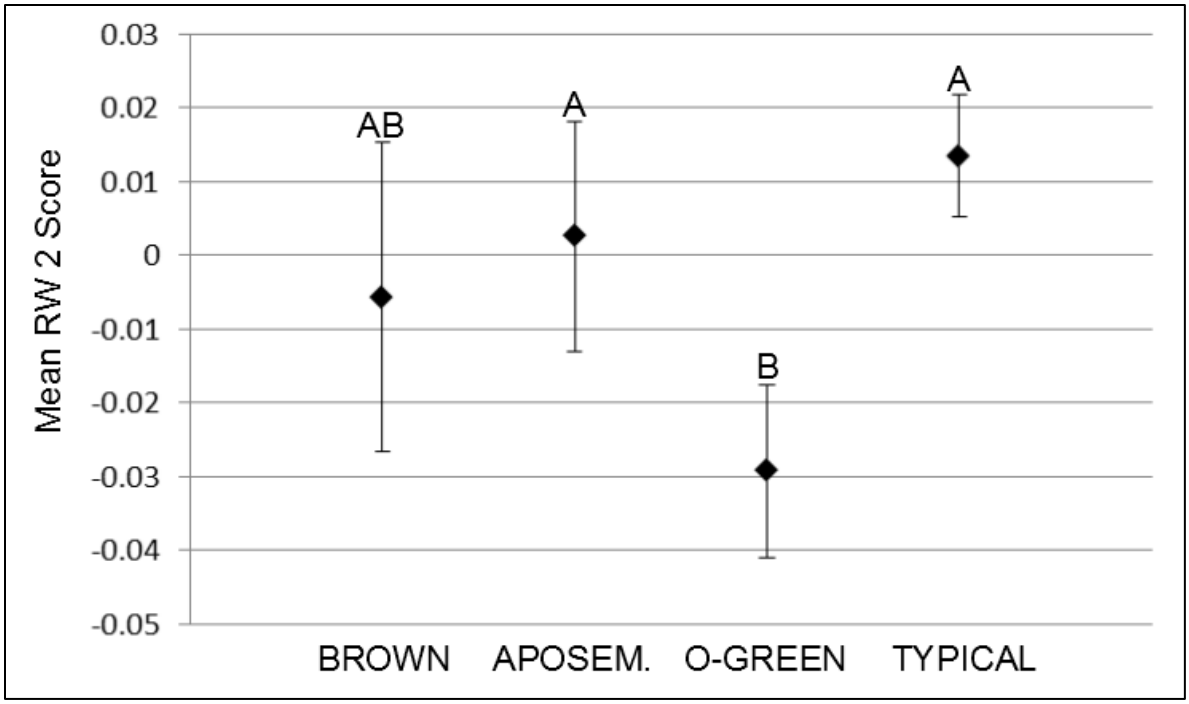

Figure 3: Mean shape (RW 2) by ecotype
Means are least squared means. Error bars represent 95% confidence intervals and capital letters are Tukey's groupings for ecotype.

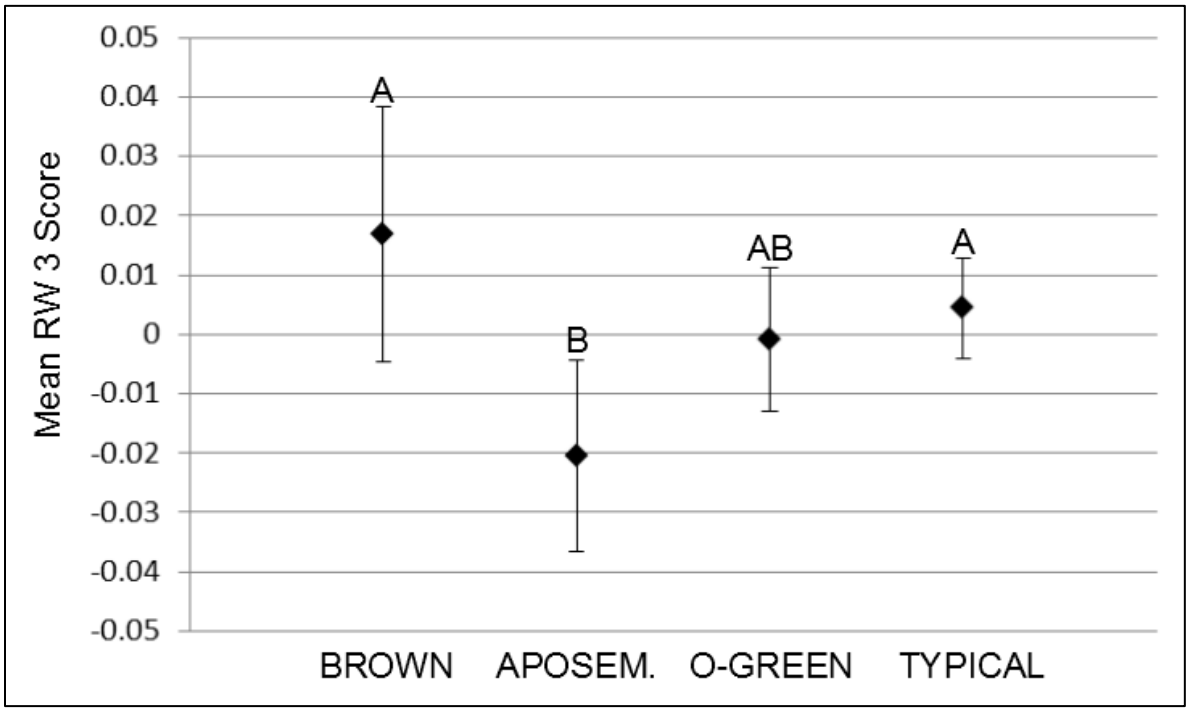

Figure 4: Mean shape (RW 3) by ecotype
Means are least squared means. Error bars represent 95% confidence intervals and capital letters are Tukey's groupings for ecotype.

Although RW 2 and RW 3 together only explained 33.48% of the observed variation, they were the only two metrics in which variation was driven by the effect of ecotype, making them appropriate for visualization of the morphospace divergence across ecotypes. I plotted the scores of RW 3 as a function of RW 2 and illustrated minimum convex polygons (MCPs) to represent the relative occupied morphospaces for cerci of each ecotype (Fig. 5). This type of visualization is interpreted by looking at the overlap in MCPs for each group. When groups are morphologically distinct, the morphospaces are expected to overlap very little or not at all; more overlap indicates a higher degree of similarity. We found that there is a high amount of variation in shape both within and across ecotypes. The typical and olive-green ecotypes display a greater amount of variation within an ecotype than the brown and aposematic ecotypes, as shown by the larger area of the MCPs in those ecotypes. The morphospaces for all four ecotypes overlap to some extent, indicating that although there may be a signal of divergence between ecotypes, there is still some similarity. I also visualized the shape of cercus at specific points at the extreme ends of the axes (Fig. 5).

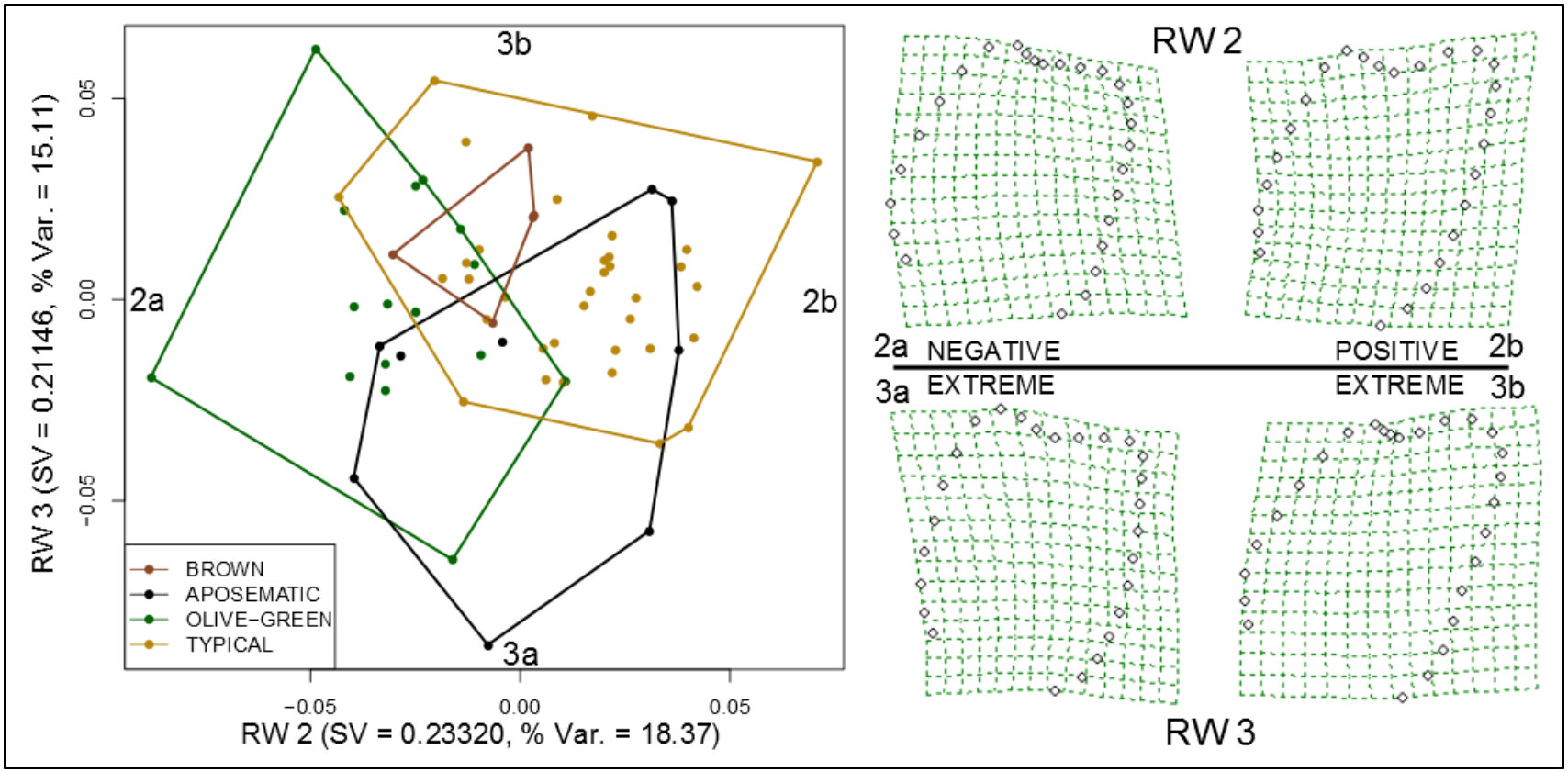

Figure 5: Relative morphospaces of male cerci

RW 3 is plotted as a function of RW 2 for each of four ecotpyes on the left. Minimum convex polygons (MCPs) denote boundaries of the occupied morphospaces. Singular value (SV) and percent of variation explained (% Var.) are shown in parentheses. Hypothetical extremes generated using TPS software for both RW axes are illustrated on the right and correspond to the shapes that would occur at the points denoted on the plot to the left. 2a and 2b represent extremes for RW 2; 3a and 3b represent extremes for RW 3.

Phylogeography

PartitionFinder recommended the use of a single partition consisting of the entire concatenated data set under the HKY + I + G model. All four runs of the Bayesian analysis converged. Although numerous trees were actually constructed as part of my analysis, I present here the haplotype tree based on the BIC best-fit partitioning scheme as recommended by PartitionFinder (Fig. 6).

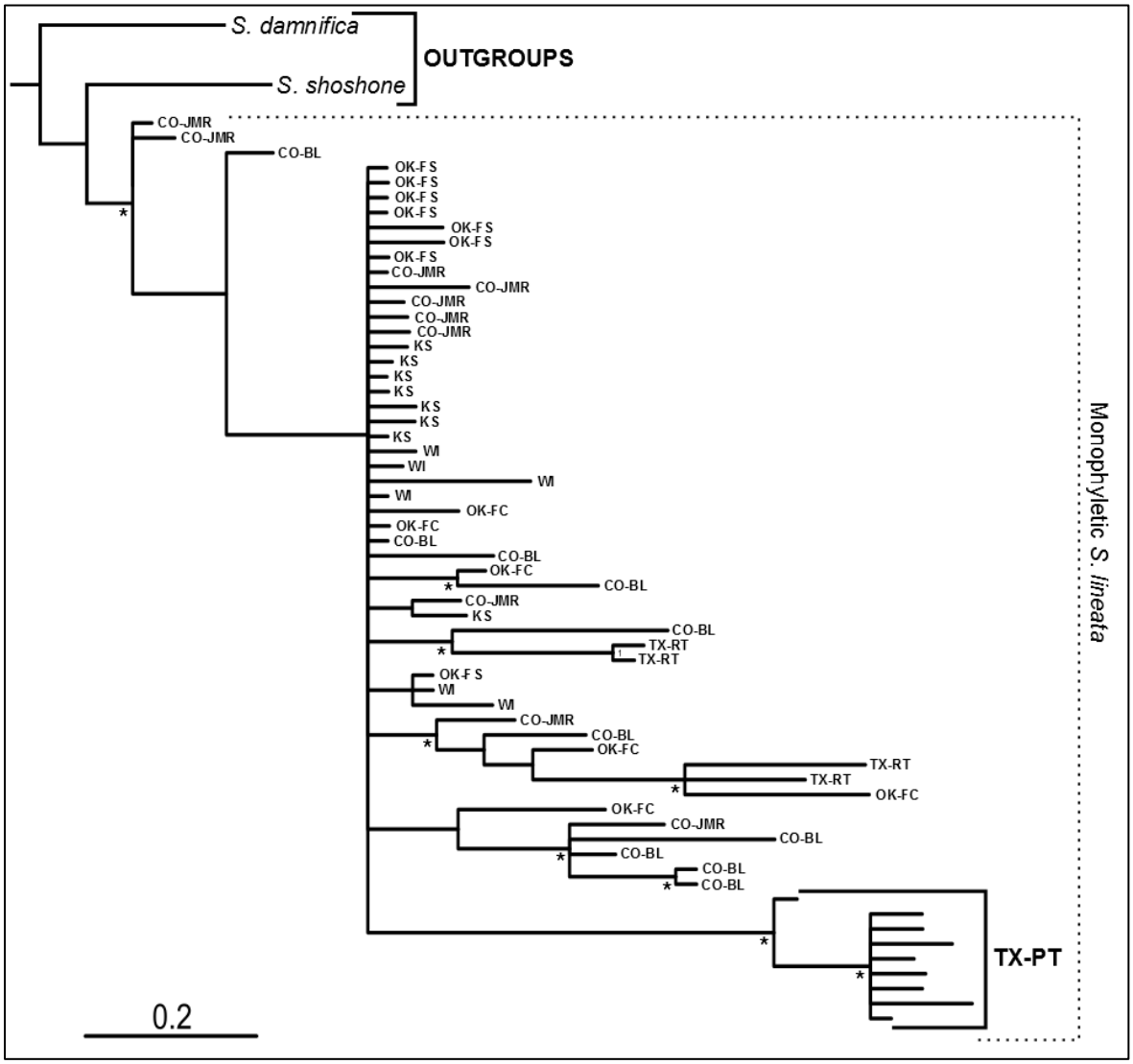

Figure 6: Bayesian haplotype tree for *S. lineata*
Terminal node labels indicate collecting locality.
*Internal nodes denoted with an asterisk show strong support (posterior probability ≥ 95%).

Removal of any one gene or any combinations of multiple genes from the analysis did not affect the topology of the tree at nodes with strong support. The analysis informed by Gblocks likewise revealed a pattern of genetic divergence congruent with the PartitionFinder-based analysis. JModelTest recommended use of a HKY + I model for COI and 12S and a GTR + I model for the 16S gene.

In both analyses, *Schistocerca lineata* was recovered as a monophyletic clade with strong nodal support, and as sister to *S. shoshone* (Thomas, 1873), which is consistent with published literature [17]. A large majority of the taxa were recovered as a large polytomy that was mostly undifferentiated internally. One notable exception was the TX-PT population, which was relatively well-differentiated from the other populations, forming a distinct monophyletic group within the polytomy, although its relationship within the rest of the clade is unclear. Neither the geography nor the assumed ecology of these populations was reflected in the topology of the tree, with the exception of the TX-PT population, representative of the aposematic ecotype.

The partial 16S gene tree including data from Dopman *et al.* (2002) shows a similar pattern to my Bayesian tree based on four mitochondrial genes (Fig. 7). Taxa representing the aposematic ecotype collected by Dopman *et al.* (2002) form a monophyletic clade with specimens from the TX-PT population. Furthermore, specimens designated as *Rubus*-feeding (a subset of the typical ecotype), which form a monophyletic group sister to the aposematic ecotype in the Dopman et al. (2002) study, are distributed throughout the polytomy with specimens from other ecotypes and without forming a distinct clade.

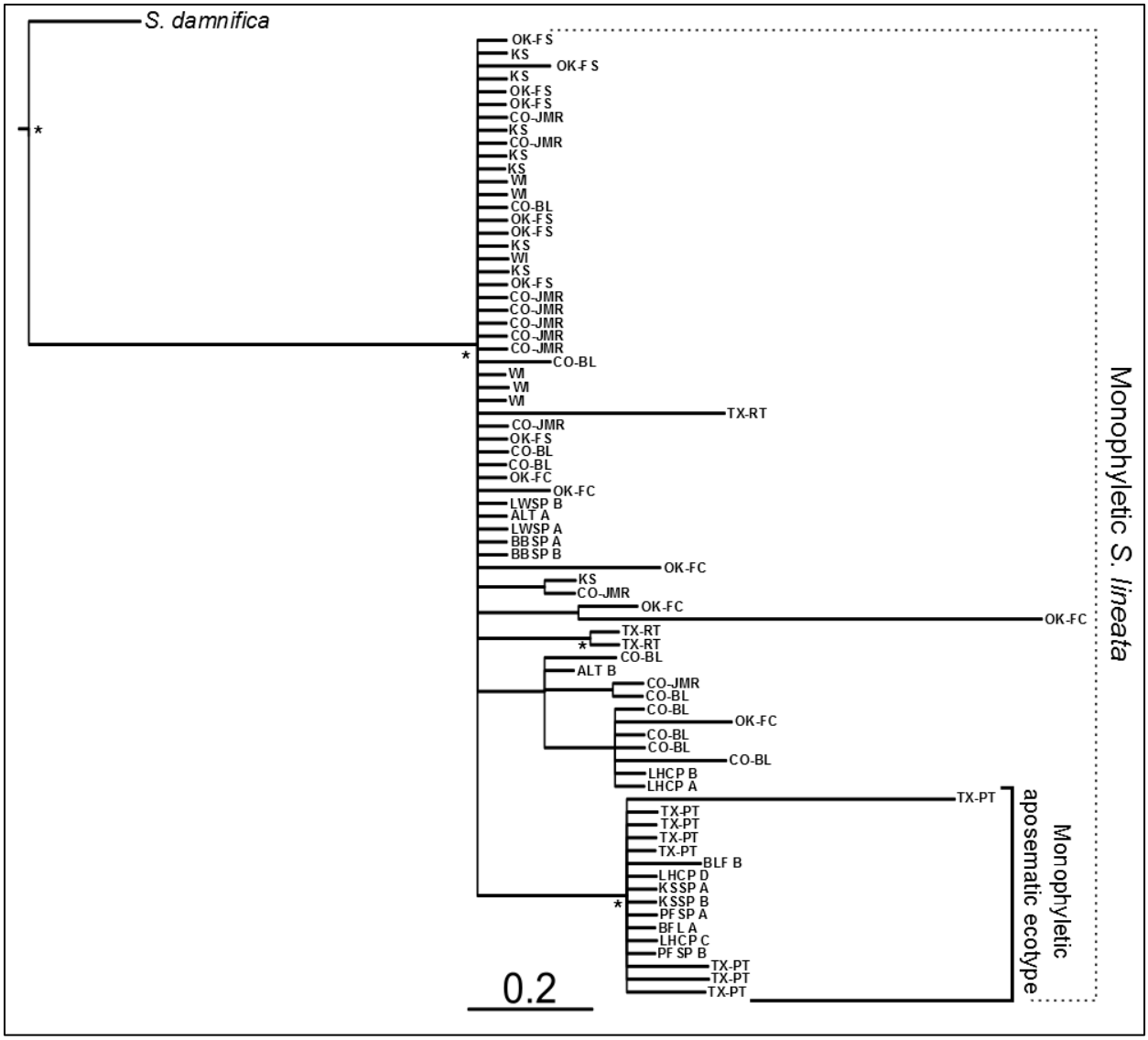

Figure 7: Bayesian 16S gene tree including data from Dopman *et al.* (2002)
Terminal node labels indicate collecting locality.
*Internal nodes denoted with an asterisk show strong support (posterior probability ≥ 95%).

Discussion

The ecotype concept implies that different ecological units within a species are following their own evolutionary trajectories [4,5,9,10], which can be reflected by divergence in ecological, morphological, or genetic traits. However, this concept is often not tested empirically, and I have shown that the ecotype concepts within *S. lineata* do not corroborate with

the data. Song (2004) originally described four ecotypes (aposematic, brown, olive-green, and typical) based on museum specimens as well as published and anecdotal ecological information on host-plant association. I find that only the aposematic ecotype (TX-PT) is genetically distinct and I find no evidence for genetic clustering of the other three ecotypes. In terms of morphology, I find no pattern supporting the distinctness of any of the four ecotypes. My study suggests that a mosaic pattern of evolution may accurately describe the patterns of character divergence from both morphological and molecular traits when the taxon of interest is undergoing the early stages of ecological speciation. I draw this conclusion from the patterns observed from three distinct character sets: (1) size, (2) shape, and (3) mitochondrial DNA and below I expand further on my findings.

Size is an Insufficient Indicator of Ecotype

My study shows that size alone is not an effective diagnostic trait for all ecotypes in *S. lineata*. Song and Wenzel (2008) compared three populations of *S. lineata* from Colorado, Kansas and Oklahoma and suggested that each population has a characteristic body size. However, my present study includes these three populations (CO-JMR, KS, and OK-FS), as well as five additional populations, and I observed that the pattern described in Song and Wenzel (2008) disappears as more data are analyzed simultaneously.

Although size may be a good character for distinguishing some ecotypes from one another, not all ecotypes have a characteristic size. The brown ecotype (WI) and the olive-green ecotype (CO-JMR, CO-BL) are both significantly smaller than the aposematic and typical ecotypes, and the brown ecotype is smaller still than the olive-green ecotype, but the aposematic ecotype (TX-PT) is not statistically different from the typical ecotype (OK-FS, TX-RT, KS, and OK-FC) (Fig. 2). There may be biological reasons as to why size is confounded as a diagnostic

character for the designation of ecotypes. For one, insect size is known to be intimately tied to the availability of suitable nutrition [46]. A variety of factors may contribute to local fluctuations in availability of suitable host plants for populations of *S. lineata*. For instance, some populations were collected during drought years, in which associated host plants may have been sparse or of poor quality. Furthermore, some populations were collected as adults in the field, whereas others were collected as nymphs and reared in the lab. It is possible that the rearing conditions affected the adult size lab-reared specimens. There may also be a latitudinal or geographic effect, as Bergmann's rule may hold true within widely distributed insect species [47]. These, and other environmental factors, could potentially influence the overall size of a population, potentially from year-to-year, and it is therefore necessary to examine traits that are independent of size to fully understand the pattern of ecotype divergence in my system.

Shape is an Insufficient Indicator of Ecotype

Similarly to size, I find that the shape of male cercus alone does not clearly delineate all ecotypes. Using the Tukey's HSD tests and minimum convex polygons (MCPs) based on RW 2 and RW 3 scores, I find that different size-independent morphological characters appear to diverge independently from one another. The brown ecotype (WI) has the narrowest shape divergence within the population, but its male cercus shape is embedded in the center of the morphospace, overlapping with all other ecotypes. This low amount of variation may be influenced by small sample size (n = 5). Within the typical ecotype (TX-RT, OK-FS, and OK-FC), I observe the largest amount of shape variation, occupying the broadest shape space. The original designation of this ecotype was not based on any particular morphological or ecological traits, which would indicate that it was an arbitrary designation for calling ecologically divergent populations as one ecotype [17]. There is moderate overlap between the typical ecotype and

olive-green ecotype, which also displays a large amount of intra-ecotype variation; however, these ecotypes appear to diverge towards opposite extremes of RW 2, and the olive-green ecotype is significantly different from the typical and aposematic ecotypes for RW 2 (Fig. 3). When examining RW 3, the aposematic ecotype emerges as a divergent group, being significantly different than both the brown and typical ecotypes (Fig. 4).

The observed pattern suggests that a considerable amount of shape variation in male cercus exists in *S. lineata* within the limit of species-specific shape. Some characters diverge in shape across ecotypes, but others are less variable. Additionally, some characters diverge differentially in different ecotypes. Song and Wenzel (2008) compared the shape of male cercus among OK-FS, KS, and CO-JMR and found that CO-JMR had a distinctly different shape, but when I expanded this study to include more ecotypes, the uniqueness of CO-JMR disappeared.

It is important to consider the function and possible selective pressures that a morphological trait faces when studying ecotype-level variation. Functionally, male cerci may play an important sensory role during copulation in *S. lineata* (Song, pers. comm.). In fact, the shape of male cerci is considered an important species diagnostic character in *Schistocerca*, which implies that there may be a strong selective pressure, possibly from sexual selection, resulting in rapid morphological divergence [17]. If male cerci were indeed under selection, I would expect to see the largest amount of shape divergence between the ecotypes that are geographically close. Two populations in Texas (TX-PT and TX-RT) are geographically close, but ecologically very divergent, representing the aposematic and typical ecotypes respectively, which may result in a reduced selective pressure on male cercus. Of course, there is a possibility that divergence or non-divergence of male cerci is simply a by-product of drift and has little to do with selection.

mtDNA is an Insufficient Indicator of Ecotype

I find that even a fairly comprehensive molecular dataset consisting of four mitochondrial loci is not capable of recovering a phylogeographic pattern of relatedness for four ecotypes of *S. lineata*. This is unusual because it implies that mtDNA markers, which are commonly used for estimating divergence at the intra-specific level, may be insufficient for this system. Alternatively, it implies that the rate of divergence of populations or ecotypes within *S. lineata* may be more rapid than that of mtDNA. One exception is the aposematic ecotype (TX-PT), which stands out as a strong monophyletic group with a relatively long branch (Figs. 6 & 7).

Dopman *et al*. (2002) studied whether two known host races of *S. lineata* in Texas (represented by TX-PT and TX-RT in my study) were genetically distinct from each other. They found that the host race associated with feeding on toxic *Ptelea trifoliata* and another feeding on non-toxic *Rubus trivialis* formed reciprocally monophyletic groups and suggested that resource-associated divergence might have played an important role in the formation of these two host races. However, *S. lineata* exists as many more host races, populations, and ecotypes, and Dopman *et al*. (2002) only included the two host races in their study along with several divergent outgroup species [11,12]. In my study, I find a congruent pattern that the *Ptelea*-associated aposematic ecotype (TX-PT) is genetically distinct, but I do not recover a monophyly of the *Rubus*-associated population (TX-RT). When including data from Dopman *et al.* (2002), I recover a polytomy consisting of all non-*Ptelea*-feeding populations, and the eight *Ptelea*-feeders from that study (representing four other collecting localities) form a monophyletic group with my aposematic TX-PT population. The result from Dopman *et al*. (2002) might have been somewhat biased because they only included two feeding groups – one of which is genetically very divergent, which by default makes the other ecotype monophyletic. We find that this

reciprocal monophyly breaks down when including specimens from other feeding groups, but that all *Ptelea*-feeding populations form a well-supported clade.

This shows that resource-associated genetic divergence clearly exists in *S. lineata*, but is only demonstrated in the *Ptelea*-associated aposematic ecotype and not for other ecotypes or host races, based on mtDNA data.

Mosaic Pattern of Character Divergence

When different characters show different patterns of divergence, mosaic evolution is said to have occurred [33]. In this study, I have shown that size, shape, and DNA are unable to consistently delimit ecotypes when each trait is examined individually, and furthermore, they show different patterns of divergence from each other. With regards to ecology, it is known that host plant association has evolved several times, including the well-characterized cases of *Ptelea*-feeding and *Rubus*-feeding populations and a suspected case of a *Tamarix*-feeding population [17,18]. In terms of adult coloration, there are at least three unique morphs (aposematic, olive-green, and brown) and one type (typical) that appears to be highly variable. When size, shape, DNA, ecology, and color, are collectively considered, it is possible to identify distinct groups within *S. lineata*, but it is clear that different characters diverge separately from each other.

The most distinct ecotype within *S. lineata* is the aposematic ecotype associated with *Ptelea trifoliata*. It has a characteristic yellow and black pattern both as adults and nymphs, and it is genetically divergent from other populations. However, in terms of body size and male cercus shape, it is not necessarily distinct from other groups. On the other hand, a *Rubus*-feeding population (TX-RT) is distinct only in terms of feeding ecology. Its color is non-descriptive, and it is not genetically distinct from other populations. It is also indistinguishable from others in

terms of body size and male cercus shape. The brown ecotype from Wisconsin is smaller than any of the populations and characteristically brown in color, but it is not distinct in terms of male cercus shape or mtDNA. The olive-green ecotype has a unique color pattern and possibly a distinct feeding ecology, but it is not genetically distinct, and it is morphologically divergent by some measures of morphology, but not by others.

Understanding how reproductive isolation (barriers to gene flow) evolves is fundamental to studying speciation [1,3,8,28]. While the best-studied systems emphasize the role of ecological divergence or geographical isolation, in reality, most species must experience a combination of selection, drift, and demographic stochasticity, which can result in numerous divergent-looking populations or ecotypes [4,5,6,7,8,9,10,29]. From my study, it is clear that there is no one path that results in an independent evolutionary trajectory. Also, my study suggests that within a widespread, highly mobile, herbivorous species with a tendency to associate with host plants, there can be multiple evolutionary trajectories varied in their positions along a speciation continuum, which can eventually evolve into reproductively isolated new lineages. For example, I believe that the *Ptelea*-associated population (TX-PT) is much further along a speciation continuum than the *Rubus*-associated population (TX-RT) and that they are headed towards two different trajectories. The rates at which different morphological, ecological, and genetic traits accumulate along these trajectories differs considerably among different populations, and there is no one clear explanation that describes the pattern, other than the fact that a mosaic pattern of character divergence is expected at this level of differentiation.

CHAPTER III: MICROBIAL COMMUNITY STRUCTURE RELFECTS FEEDING ECOLOGY

Introduction

The spotted bird grasshopper *Schistocerca lineata* Scudder, 1899 (Orthoptera: Acrididae) is a widely distributed North American species that occurs in highly localized and potentially isolated populations that are often associated with different host plants [17,18,24]. Although there is a diversity of color morphs and feeding habits within the species, there are only two feeding groups with well-characterized ecologies [19,20]. Both found in Texas, the first feeding group is associated with dewberry, *Rubus trivialis* Michx., and is tan in color, while the second group feeds almost exclusively on the toxic wafer ash, *Ptelea trifoliata* (L.), and displays density-dependent aposematism and derives chemical defense from its host plant [20,21,22]. Nymphs of *Ptelea*-feeders have been shown to be highly preferential in feeding habit and consume wafer ash readily, whereas *Rubus*-feeding nymphs will refuse wafer ash and starve if not provided an alternative host option [20]. This suggests that *Ptelea*-feeders may be equipped in some way to deal with the *P. trifoliata* toxicity, whereas *Rubus*-feeders are not.

A potential explanation for the *Ptelea*-feeders' ability to deal with a toxic host plant may involve a tri-trophic interaction. Bacterial endosymbionts have been shown to help herbivorous insects withstand gut-borne pathogens and may promote rapid evolution and adaptation in their hosts [25,26,48,49]. They are also known to contribute to nutrition for insects on suboptimal diets [25,26]. In the desert locust *Schistocerca gregaria*, the gut microbial community is dominated by a relatively low number of species acquired from the diet [50]. These microbes metabolize secondary plant chemicals that might otherwise be detrimental to the host [48]. Considering that *P. trifoliata* is known to be cytotoxic, it may be that *Ptelea*-feeders harbor a

microbial community different from that of *Rubus*-feeders (who may not need to metabolize particular plant secondary chemicals), enabling them to associate with such a toxic host plant [51]. Here, using early instar nymphs, which are most strongly-associated with their host plants, I assess the microbial community structure of eight populations of *S. lineata*, and relate this structure back to the known ecology of those grasshoppers. This study makes use of a next-generation sequencing (NGS) technique called 454 pyrosequencing and subsequent bioinformatics to generate absolute and relative bacterial abundance data for 79 individual grasshoppers. I hypothesize that populations of the same ecotype will cluster statistically based on the relative composition of their respective microbial communities.

Methods

Taxon Sampling

I collected 10 first, second, and third instar nymphs from 5 populations of *Ptelea*-feeders and 3 populations of *Rubus*-feeders across their native range in Texas (Table 5). For one population (LCR), only nine individuals were available. These specimens were collected directly into 100% ethanol and complete genomic DNA was extracted using the QIAGEN DNeasy Kit. Because this kit is non-specific, this process also extracts DNA from all bacteria in the sample. Although I originally planned to target the insects' midguts, the early instar nymphs proved too small to dissect efficiently and extract a high enough concentration of quality DNA. Instead, we used DNA from the entire body of the grasshopper, except for the hind legs which were removed as backup genetic material.

Table 5: Collection summary and description of nymphal habitat

Population (abbr.)	Host Plant	Habitat
Lake Whitney (LW)	*Rubus trivialis*	shrubby lakeshore, minimal canopy
College Station (CS)	*Rubus trivialis*	grass and shrubs, arid, no canopy
Altair (AT)	*Rubus trivialis*	dense *R. trivialis*, arid, no canopy
Balcones Canyonlands Doeskin Tract (BCD)	*Ptelea trifoliata*	elevated hillside, near water, no canopy
Brackenridge Field Lab. (BFL)	*Ptelea trifoliata*	woody area, some canopy
McKinney Falls (MKF)	*Ptelea trifoliata*	woody area, near water, thick canopy
Lower Colorado River (LCR)	*Ptelea trifoliata*	woody area, near water, thick canopy
Bamberger Nature Park (BNP)	*Ptelea trifoliata*	woody area, some canopy

The whole nymph was ground with a motorized pestle and tissue lysed overnight at 56ºC. After extraction, DNA concentrations were quantified using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc.). The set of 20 µL aliquots were then shipped in dry ice overnight to Research and Testing Labs, Lubbock, Texas.

Molecular Character Sampling

The bacterial 16S genes for all bacteria in an individual were amplified using 454 pyrosequencing. Each sample was fixed with a unique NGS sequencing barcode, which was used to identify the individual grasshopper of origin for each bacterial sequence. These sequences were then denoised and assembled *de novo*. Denoising included trimming of low quality ends from each read or trimming of entire reads that did not meet quality standards, removal of sequences that failed to cluster with other reads, and elimination of chimeric sequences. Clusters of reads known as operational taxonomic units (OTUs) were then assigned a likely species of origin identified using BLAST, providing presence-absence data for every bacterial species across the dataset. It was also possible to infer the relative abundances of each bacterial species in each grasshopper by estimating from the number of reads per OTU. I checked for sampling bias with regards to concentration of genomic DNA by carrying out an ANOVA and a Tukey's HSD test in JMP Pro 10.0 using population as the categorical predictor of DNA concentration. I also ran a regression to test for correlation of microbial diversity (I transformed the raw bacterial generic richness for each individual using the Shannon-Weiner Index) with initial DNA concentration.

Hierarchical Clustering Analyses

I carried out a hierarchical clustering analysis (HCA) in R 2.15.1 based on the mean relative abundances of the 24 most abundant bacterial genera, as well as based on the entire

dataset (98 bacterial genera) and compared the clustering pattern to the ecology of the two feeding groups [19]. I carried out the HCA at the generic level because it allowed me to account for a much larger portion of the dataset then if I were to analyze at the specific level.

Results

There were some significant differences in the initial concentrations of DNA in the raw genomic extracts across populations (Fig. 8). Some samples did not amplify and/or sequence efficiently and were dropped from the analysis (n = 72).

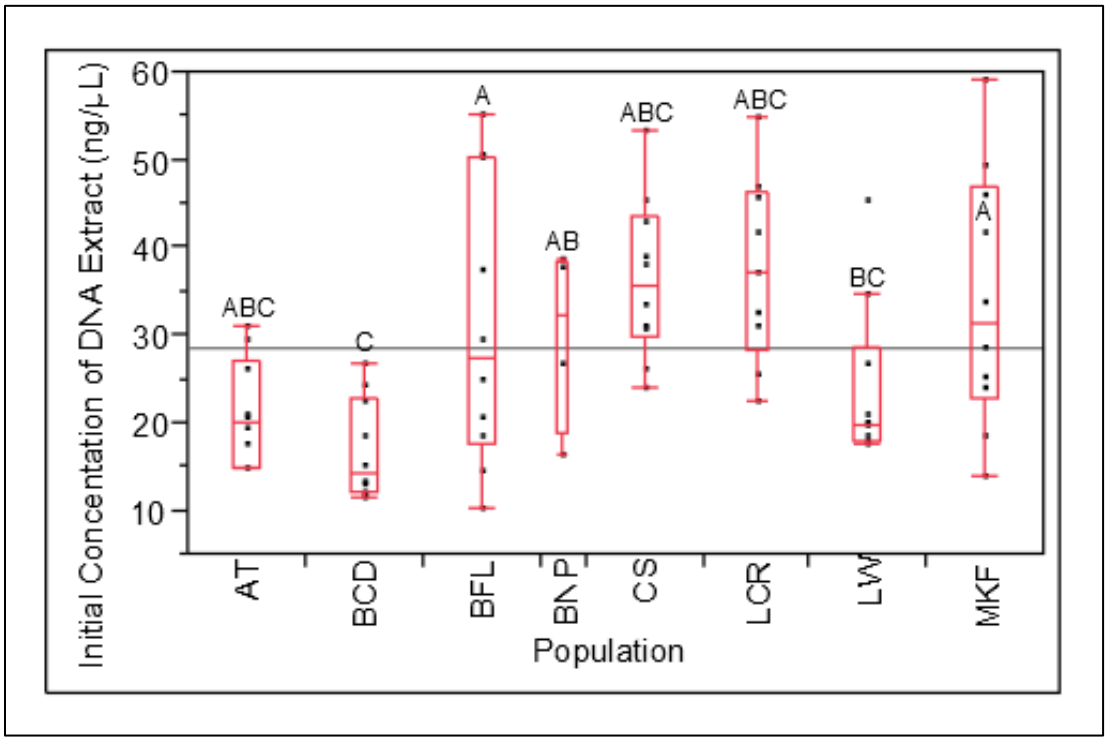

Figure 8: Mean concentrations of DNA extract for eight populations of *S. lineata*
ANOVA results: $p = 0.0003$, $F_{7,65} = 4.6860$. Capital letters are Tukey's groupings.

Although most of the populations showed relatively similar mean concentrations, the mean concentration for the BCD population was significantly lower than those of the BFL and MKF populations. These means were associated with a relatively large variance, and my regression of

microbial diversity by DNA concentration showed a very weak correlation (adj. $R^2 = 0.041667$), providing evidence for a negligible effect of initial DNA concentration (Fig. 9).

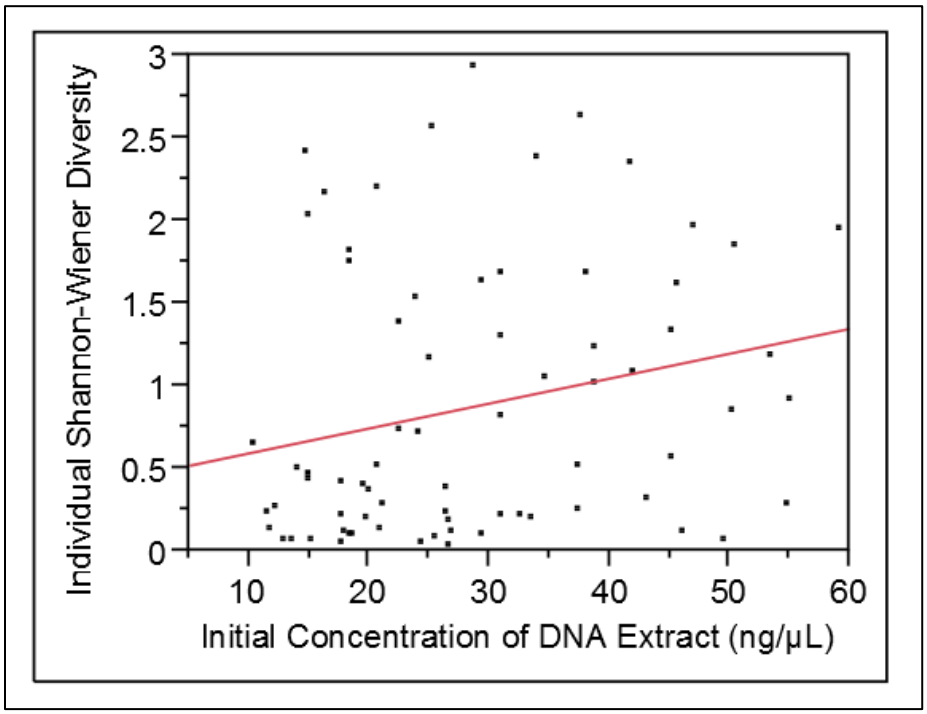

Figure 9: Correlation of diversity by initial concentration of DNA extract
Linear regression: $y = 0.4395748 + 0.0150857x$; adj. $R^2 = 0.041667$.
ANOVA results: $p = 0.0459$, $F_{1,71} = 4.1305$**.**

Hierarchical Clustering Analyses

My HCA based on relative bacterial abundance of the 24 most abundant bacterial genera revealed a clustering pattern that nearly reflected the ecology of the host grasshoppers, with a single exception; BCD, a *Ptelea*-feeding population, clustered with the three *Rubus*-feeding populations (Fig. 10). These 24 most abundant genera accounted for roughly 97% of the total bacterial abundance across all populations. One can also see from the heat map, that bacteria of the genus *Methylobacterium* dominate the microbial communities in all eight populations, contributing a mean of 64.1% of the total abundance across the dataset.

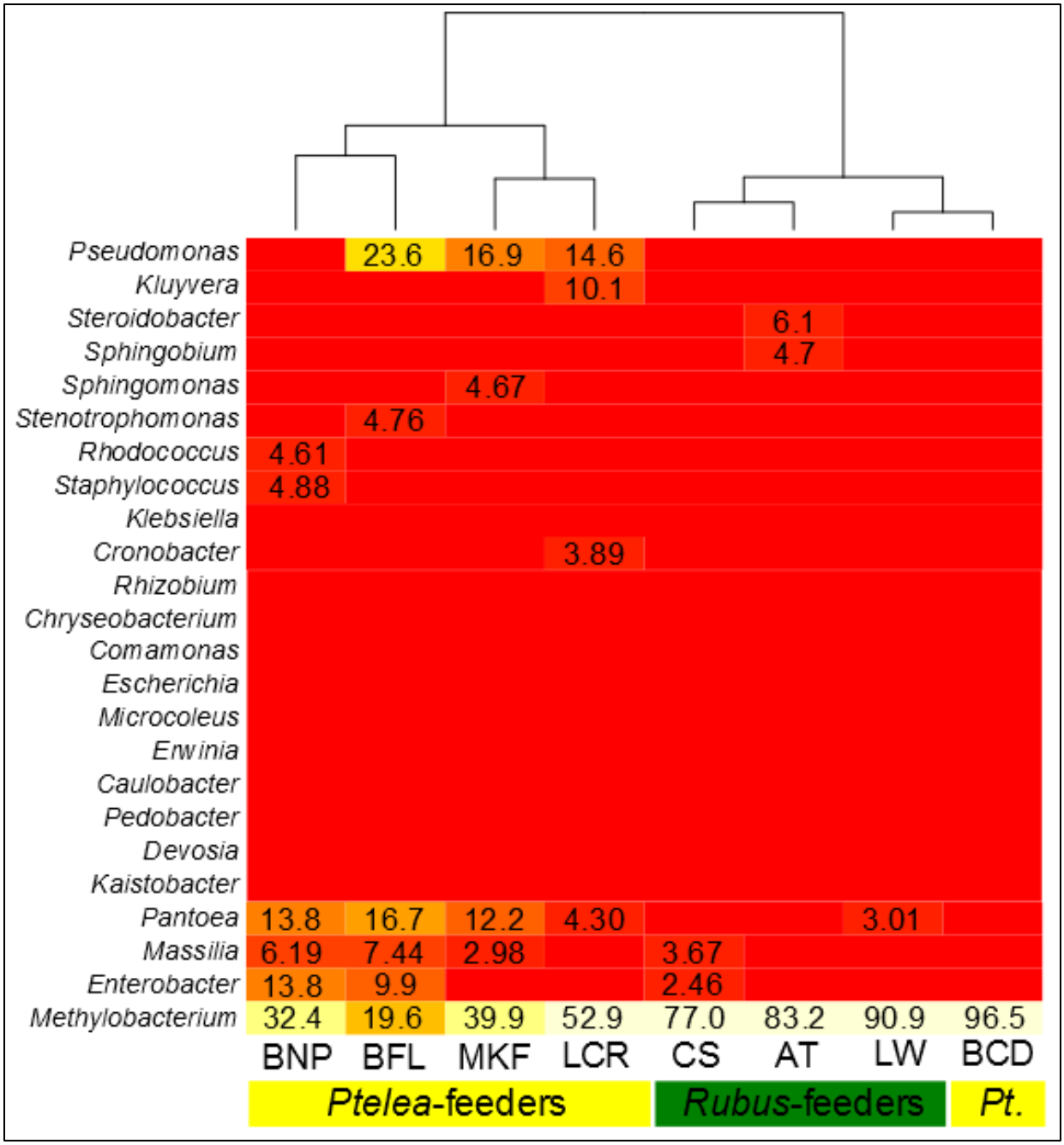

Figure 10: Clustering dendrogram (HCA)
Illustrates similarity of the bacterial community structure across populations, based on the 24 most abundant bacterial genera. Lighter colors indicate higher relative abundance. Percentages in each box indicate the mean relative contributions of those genera to the bacterial community in each population. Ecotype for each population is shown at the bottom.

In four of the *Ptelea*-feeding populations, bacteria of the *Pseudomonas*, *Pantoea*, *Massilia*, and

Enterobacter genera are relatively abundant, likely causing the strong clustering pattern seen in

the dendrogram, although *Kluyvera* has relatively high abundance in one of these populations, and some genera are not as abundant in some populations as they are in others. Although I do not present it here, the HCA including all bacterial genera yielded the same clustering topography as the analysis run with only the 24 most abundant bacteria.

Discussion

This study provides a unique look into the potential role of bacteria in ecological divergence in a host-associated grasshopper. I found that there could be vast differences in microbial community structure between populations within an ecotype. Assuming that bacterial identification of OTUs was reliable, our dataset provides a unique opportunity to explore life histories of the dominant bacterial genera and explore other aspects of ecology with regards to the microbial community structure observed in this system.

Microbial Community Reflects Ecology

Nymphs of *S. lineata* likely hatch in an area of soil near an appropriate host plant since females of host-associated species tend to oviposit in close association with that host plant [24,52]. Thus, it stands to reason that nymphs quickly find their way to their first meal, and remain in close proximity to the host plant throughout their early life. The bacterial load observed in these populations can likely be explained by the unique life history traits of this species. However, it is crucial to understand the bacteria themselves, and relate this information back to the ecology of the grasshoppers to fully understand the data.

Methylobacterium is a genus of nitrogen-fixing bacteria commonly associated with soil, leaves, and other plant parts [53]. Because this group was found to be cosmopolitan and to be the dominant microbial genus across all populations, it is likely that the grasshoppers are

exposed to *Methylobacterium* spp. throughout their entire lives, which is concordant with the life cycle of the species.

Pseudomonas, *Massilia*, *Pantoea*, and *Enterobacter* are the four microbial genera that seem to drive the clustering of the BNP, BFL, MKF, and LCR populations in my HCA (Fig. 10). *Pseudomonas* is a diverse genus of aerobic gammaproteobacteria that includes members species ranging from human and plant pathogens to soil flora [54]. Like *Methylobacterium* spp., *S. lineata* nymphs are probably exposed to *Pseudomonas* spp. early in life in the soil, and probably continue to acquire the bacteria from association with the host plant. *Massilia* is a similarly diverse genus, with samples being isolated from human patients, soil, and the air [55,56,57,58].

Of particular interest are the *Pantoea* and *Enterobacter* bacterial genera, both of which belong to the Enterobacteriaceae family. These genera have been identified as part of the natural microbial community for a number of insects, and some species have been shown to confer benefits to the host grasshopper [25,48,50]. Microbial metabolism of secondary plant chemicals by Enterobacteriaceae in the locust gut has been shown produce phenolics useful to the locust host [48]. *Pantoea aggomerans*, in particular, which accounts for 3.01% of the bacterial abundance across the dataset and 4.52% of the bacterial community in *Ptelea*-feeders, has been shown to produce an antifungal compound in grasshopper guts [48]. This collectively suggests that *Ptelea*-feeders, whose microbial communities harbor relatively high abundances of these genera, may receive some benefit from these bacteria, perhaps through metabolism of *P. trifoliata* cytotoxins. Again, it is most likely that these microbes are acquired from the diet, which raises an interesting question – If *Ptelea*-feeding *S. lineata* populations draw their ability to metabolize plant toxins from diet-acquired bacteria, then why do *Rubus*-feeders outright reject *P. trifoliata* as a suitable host plant?

Misunderstood Interactions in Complex Communities

Grasshoppers rely on visual and chemical cues to identify suitable host plants [24,52]. However, most polyphagous species will readily consume other, less-preferred plants rather than starve. This is not the case when presenting *Rubus*-feeders with *P. trifoliata*, who will die without even attempting to consume the plant [20]. A potential explanation for *Rubus*-feeders' rejection of *P. trifoliata* could be that they are cuing in on the plant's toxicity and avoiding it, whereas *Ptelea*-feeders might be attracted to that same chemical cue. Interestingly, my data suggest that grasshoppers of both ecotypes acquire their microbial communities from association with their host plants. If some component of microbial diversity confers the ability to metabolize *P. trifoliata*, then *Rubus*-feeders should theoretically be able to consume *P. trifoliata* and, in doing so, acquire the necessary microbial symbionts needed to deal with the plant's toxicity. This reveals an important disconnect in our understanding of the interactions of plant toxicity and microbial metabolism because if diet-acquired bacteria allow for processing of plant toxins, then any insect should be able to consume that plant, essentially negating the benefit of toxicity to the plant. Furthermore, if non-associated insects could possibly herbivorize a toxic plant but do not do so because of some cue, then what could induce an initial host plant shift to that plant?

Another wrench in the proverbial cogs is represented by my population of *Ptelea*-feeders collected from Balcones Canyonlands – Doeskin Tract (BCD). This population was one of the largest I observed in the field, with hundreds of nymphs on dozens of *P. trifoliata* bushes along an upland stream. This population harbored relatively low abundances of bacterial genera that were found in other *Ptelea*-feeding populations [48,50]. Additionally, the *Pantoea* and *Enterobacter* genera, which were the most likely candidates for a symbiosis in this system, are in relatively low abundance compared to the other *Ptelea*-feeding populations (Fig. 10). This

suggests then that the microbial community structure may be influenced by factors other than feeding preference. Notice that for all *Ptelea*-feeding populations other than BCD, there is some plant canopy structure, whereas the lack of canopy observed at BCD reflects a local plant assemblage more similar to those of the *R. trivialis* habitats (Table 5). It is possible that the local microflora available to the grasshoppers is influenced by plant community structure or that the BCD population is otherwise ecologically different from the other *Ptelea*-feeders in some unidentified way. Alternatively, it may be that multiple bacteria play similar roles in different populations of grasshoppers feeding on the same host plant. For a more robust understanding, bacteria should be identified to species, and their origins in the grasshopper body as well as their function within the animal should be fully explored. Furthermore, examination of the biochemistry and metabolism of *P. trifoliata* after consumption by the grasshopper will be crucial to understanding the tri-trophic interactions in this system.

CHAPTER IV: GENERAL DISCUSSION AND CONCLUDING REMARKS

From the literature and from my two studies, I can conclude that the *Ptelea*-feeding, aposematic ecotype of *Schistocerca lineata* represents a genetically and ecologically distinct lineage from the rest of the species. I identified a mosaic pattern of character evolution across size-dependent and size-independent measures of morphology, indicating that morphology may be unreliable for demarcation of ecotype. Genetic data provided a clearer, more robust indication of ecotype, suggesting that morphologically-defined or ecologically-defined ecotypes in other taxa should be evaluated from a phylogeographic perspective, especially since we interpret the term "ecotype" as reflecting distinct evolutionary lineages. While ecological divergence may eventually lead to reproductive isolation in some taxa, ecological traits found in a population might also be ephemeral, having no genomic consequence.

I also found evidence that microbial community composition may differ across ecotypes and that microbial load is likely acquired from the diet. Although I only examined a subset of populations of *S. lineata*, the genetically distinct *Ptelea*-feeding ecotype harbored a unique and diverse microbial community that may or may not play a role in their divergence. Microbial diversity across the remainder of the species' range should also be assessed to see if this pattern truly reflects the ecology of the grasshopper or if it is a function of more complex community-wide interactions. With future work, I will illuminate the tri-trophic interactions between the grasshoppers, their host plants, and the diet-acquired bacteria, and to understand the microevolutionary consequences of these interactions in ecologically divergent species.

APPENDIX A: NOVEL PCR PRIMERS

This appendix includes information about novel PCR primers that were created to be species-specific to Schistocerca lineata for the purposes of amplification of the COI gene. The primers' names contain an "F" or an "R" to signify forward and reverse primers respectively. Primer sequences and annealing temperatures are also provided.

Name	Sequence	Anneal. Temp. (°C)
SlinF-192	CCT AAA ATT CAG CCA TCT TAC CGC	56.2
SlinF-737	TAT GAT CTG TAG CTA TTA CAG CCC	53.7
SlinF-1104	CAT CAG CAA CAA TAA TTA TTG CCG	53.2
SlinF-1327	ATT ATC TAT AGG AGC AGT ATT CGC	51.8
SlinR-leu	TTA AAT CTA CTG CAC TAA TCT GCC	52.8
SlinR-1073	TGC TCG TGT GTC AAC ATC TAT TCC	57.0
SlinR-1348	TGA ATA ACA CCT CCT ATA ATT GCG	52.8
SlinR-544	AAC TGT TCA TCC TGT ACC AGC ACC	59.0

APPENDIX B: CONCENTRATIONS OF GENOMIC DNA FOR PYROSEQUENCING

Research and Testing Labs, Lubbock, Texas, recommends DNA concentrations of 20 ng/μL or greater for reliable amplification and sequencing of the bacterial 16S gene. This appendix lists the samples (listed by population and numbered individually) that I submitted for my bacterial 16S assay and the initial concentrations of DNA for each sample.

Sample ID	(ng/μL)	Sample ID	(ng/μL)	Sample ID	(ng/μL)
AT 1	29.4	BFL 8	18.3	LCR 5	32.40
AT 2	14.8	BFL 9	55	LCR 6	45.60
AT 3	17.5	BFL 10	50.4	LCR 7	41.60
AT 4	21	LW 1	17.5	LCR 8	47.00
AT 5	26.2	LW 2	19.6	LCR 9	54.80
AT 6	20.6	LW 3	20	MKF 1	18.40
AT 7	14.8	LW 4	20.8	MKF 2	28.60
AT 8	14.8	LW 5	34.6	MKF 3	25.20
AT 9	31	LW 6	26.6	MKF 4	46.00
AT 10	19.4	LW 7	17.8	MKF 5	41.80
CS 1	53.4	LW 8	18.5	MKF 6	14.00
CS 2	38	LW 9	17.7	MKF 7	24.00
CS 3	45.2	LW 10	45.2	MKF 8	33.80
CS 4	43	BNP 1	39	MKF 9	49.40
CS 5	38.8	BNP 2	26.6	MKF 10	59.00
CS 6	30.8	BNP 3	25.6	BCD 1	15.00
CS 7	31	BNP 4	18.7	BCD 2	11.70
CS 8	23.8	BNP 5	22.8	BCD 3	11.50
CS 9	26.2	BNP 6	19.3	BCD 4	18.40
CS 10	33.4	BNP 7	19	BCD 5	26.80
BFL 1	37.4	BNP 8	16.2	BCD 6	12.00
BFL 2	14.6	BNP 9	37.6	BCD 7	22.40
BFL 3	20.6	BNP 10	38.6	BCD 8	13.40
BFL 4	50.2	LCR 1	25.4	BCD 9	12.90
BFL 5	25	LCR 2	22.4	BCD 10	24.20
BFL 6	10.2	LCR 3	31		
BFL 7	29.4	LCR 4	37.2		

LIST OF REFERENCES

1. Wiens JJ (2004) What is speciation and how should we study it? *Am. Nat.* 163(6): 914–923.

2. Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* 18: 375-402.

3. Coyne JA, Orr HA (2004) *Speciation.* Sinauer, Sunderland, MA.

4. Powell THQ, Hood GR, Murphy MO, Heilveil JS, Berlocher SH, *et al.* (2013) Genetic divergence along the speciation continuum: The transition from host race to species in *Rhagoletis* (Diptera: Tephritidae). *Evolution* 67: 2561-2576.

5. Peccoud J, Ollivier A, Plantegenest M, Simon J-C (2009) A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *P. Natl. Acad. Sci. USA* 106: 7495-7500.

6. Mani GS, Clarke BC (1990) Mutation order - A major stochastic process in evolution. *P. Roy. Soc. B-Biol. Sci.* 240: 29-37.

7. Ritchie MG (2007) Sexual selection and speciation. *Annu. Rev. Ecol. Evol.* S. 38: 79-102.

8. Rundle HD, Nosil P (2005) Ecological speciation. *Ecol. Lett.* 8: 336-352.

9. Nosil P, Gompert Z, Farkas TE, Comeault AA, Feder JL, *et al.* (2012) Genomic consequences of multiple speciation processes in a stick insect. *P. Roy. Soc. B-Biol. Sci.* 279: 5058-5065.

10. Funk DJ, Egan SP, Nosil P (2011) Isolation by adaptation in *Neochlamisus* leaf beetles: host-related selection promotes neutral genomic divergence. *Mol. Ecol.* 20: 4671-4682.

11. Funk DJ, Nosil P (2008) Comparative Analyses of Ecological Speciation. In: *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects* (Tilmon KJ, ed.). University of California Press, Berkeley and Los Angeles, CA.

12. Funk DJ (2010) Does strong selection promote host specialisation and ecological speciation in insect herbivores? Evidence from *Neochlamisus* leaf beetles. *Ecol. Entomol.* 35(1): 41-53.

13. Nosil P, Crespi BJ (2006) Ecological divergence promotes the evolution of cryptic reproductive isolation. *P. Roy. Soc. B-Biol. Sci.* 273(1589): 991-997.

14. Nosil P, Crespi BJ (2006) Experimental evidence that predation promotes divergence in adaptive radiation. *P. Natl. Acad. Sci. USA* 103(24): 9090-9095.

15. Nosil P (2007) Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. *Am. Nat.* 169(2): 151-162.

16. Nosil P, Sandoval CP (2008) Ecological niche dimensionality and the evolutionary diversification of stick insects. *PLoS ONE* 3(4): e1907.

17. Song H (2004) Revision of the Alutacea Group of genus *Schistocerca* (Orthoptera: Acrididae: Cyrtacanthacridinae). *Ann. Entomol. Soc. Am.* 97: 420-436.

18. Song H, Wenzel JW (2008) Mosaic pattern of genital divergence in three populations of *Schistocerca lineata* Scudder, 1899 (Orthoptera: Acrididae: Cyrtacanthacridinae). *Biol. J. Linn. Soc.* 94: 289-301.

19. Dopman EB, Sword GA, Hillis DM (2002) The importance of the ontogenetic niche in resource-associated divergence: evidence from a generalist grasshopper. *Evolution* 56: 731-740.

20. Sword GA, Dopman EB (1999) Developmental specialization and geographic structure of host plant use in a polyphagous grasshopper, *Schistocerca emarginata* (=*lineata*) (Orthoptera: Acrididae). *Oecologia* 120: 437-445.

21. Sword GA (2001) Tasty on the outside, but toxic in the middle: Grasshopper regurgitation and host plant-mediated toxicity to a vertebrate predator. *Oecologia* 128: 416-421.

22. Sword GA (1999) Density-dependent warning coloration. *Nature* 397: 217.

23. Christensen EM (1962) The rate of naturalization of *Tamarix* in Utah. *Am. Midl. Nat.* 68: 51-57.

24. Hubbell TH (1960) The sibling species of the Alutacea Group of the bird-locust genus *Schistocerca* (Orthoptera, Acrididae, Cyrtacanthacridinae). *Misc. Publ. Mus. Zool. Univ. Mich.* 116: 1Ð91.

25. Dillon RJ, Dillon VM (2004) The Gut Bacteria of Insects: Nonpathogenic Interactions. *Annu. Rev. Entomol.* 49: 71-92.

26. Moran NA (2006) Symbiosis. *Curr. Biol.* 16: R866-R871.

27. Brucker RM, Bordenstein SR (2013) The hologenomic basis of speciation: Gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science* 341: 667-669.

28. Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* 24(3): 145-156.

29. Turner TL, M.W. H, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* 3: e285.

30. Qiang J, Zhexi L, Shu-an J (1999) A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. *Nature* 398: 326-330.

31. Xu X, Makovicky PJ, Wang X, Norell MA, You H (2002) A ceratopsian dinosaur from China and the early evolution of Ceratopsia. *Nature* 416: 314-317.

32. Rae TC (1999) Mosaic evolution in the origin of the Hominoidea. *Folia Primatol.* 70: 125-135.

33. Carroll RL (1997) *Patterns and processes of vertebrate evolution.* Cambridge University Press.

34. Iwaniuk AN, Dean KM, Nelson JE (2004) A mosaic pattern characterizes the evolution of the avian brain. *P. Roy. Soc. B-Biol Sci.* 271(4): 148-151.

35. Knoll AH, Niklas KJ, Gensel PG, Tiffney BH (1984) Character diversification and patterns of evolution in early vascular plants. *Paleobiology* 10(1): 34-47.

36. Rohlf FJ (2010) Relative warps. Ecology and Evolution, SUNY at Stony Brook.

37. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, *et al.* (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28(10): 2731-2739.

38. Lanfear R, Calcott B, Ho SWY, Guindon S (2012) PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29(6): 1695-1701.

39. Stocsits RR, Letsch H, Hertel J, Misof B, Stadler PF (2009) Accurate and efficient reconstruction of deep phylogenies from structured RNAs. *Nucl. Acids Res.* 37(18): 6184-6193.

40. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12): 1572-1574.

41. Pfeiffer W, Stamatakis A (2010) Hybrid MPI/Pthreads parallelization of the RAxML phylogenetics code. *IEEE International Symposium on Parallel & Distributed Processing, Workshops and PhD Forum (IPDPSW)* 1(8): 19-23.

42. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucl. Acids Res.* 36: 465-469.

43. Dereeper A, Audic S, Claverie JM, Blanc G. (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol. Biol.* 10: 8.

44. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* 9(8): 772.

45. Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* 52: 696-704.

46. Mirth CK, Riddiford LM (2007) Size assessment and growth control: How adult size is determined in insects. *BioEssays* 29(4): 344–355.

47. Shelomi M (2012) Where are we now? Bergmann's rule sensu lato in insects. *Am. Nat.* 180(4): 511-519.

48. Dillon RJ, Charnley AK (2002) Mutualism between the desert locust *Schistocerca gregaria* and its gut microbiota. *Res. Microbiol.* 153: 503-509.

49. Hansen AK, Moran NA (2013) The impact of microbial symbionts on host plant utilization by herbivorous insects. *Mol. Ecol.* doi: 10.1111/mec.12421.

50. Dillon RJ, Vennard CT, Buckling A, Charnley AK (2005) Diversity of locust gut bacteria protects against pathogen invasion. *Ecol. Lett.* 8: 1291-1298.

51. Petit PG, Montagu M, Tremouillaux GJ, Chenieux JC, Rideau M (1993) *Ptelea trifoliata (quinine tree, hop tree): in vitro culture and the production of alkaloids and medicinal compounds.* In: Biotechnology in agriculture and forestry Vol. 21: Medicinal and aromatic plants (Bajaj YPS, ed.). Springer, Berlin.

52. Chapman RF, Joern A (eds.) (1990) *Biology of grasshoppers.* Wiley, New York, NY.

53. Lidstrom ME, Chistoserdova L (2002) Plants in the pink: Cytokinin production by *Methylobacterium*. *J. Bacteriol.* 184(7): 1818.

54. Madigan M, Martinko J (eds.) (2005) *Brock Biology of Microorganisms* (11th ed.). Prentice Hall, NJ.

55. La Sola B, Birtles RJ, Mallet MN, Raoult D (1998) *Massilia timonae* gen. nov., sp. nov., isolated from blood of an immunocompromised patient with cerebellar lesions. *J. Clin. Microbiol.* 36(10): 2847-2852.

56. Zhang YQ, Li WJ, Zhang KY, Tian XP, Jiang Y, *et al.* (2006) *Massilia dura* sp. nov., *Massilia albidiflava* sp. nov., *Massilia plicata* sp. nov. and *Massilia lutea* sp. nov., isolated from soils in China. *Int. J. Syst. Evol. Microbiol.* 256: 459-63.

57. Zul D, Wanner G, Overmann J (2008) *Massilia brevitalea* sp. nov., a novel betaproteobacterium isolated from lysimeter soil. *Int. J. Syst. Evol. Microbiol.* 58: 1245-51.

58. Weon HY, Kim BY, Son JA, Jang HB, Hong SK, *et al.* (2008) *Massilia aerilata* sp. nov., isolated from an air sample. *Int. J. Syst. Evol. Microbiol.* 58: 1422-1425.