HOST ASSOCIATED GENETIC DIVERGENCE AND SEXUAL ISOLATION IN THE GRASSHOPPER *HESPEROTETTIX VIRIDIS* (ORTHOPTERA : ACRIDIDAE)

by

TONY GRACE

M.S., Indian Agricultural Research Institute, New Delhi India, 2000

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

GENETICS
COLLEGE OF AGRIULTURE

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

ABSTRACT

Understanding evolutionary processes that structure genetic variation associated with lineage diversification and speciation is a central problem. Shifts in host plant use by insect herbivores and subsequent divergence in mating signals can lead to sexual isolation, especially when selection for specialization on different hosts restricts gene flow among populations. The grasshopper *Hesperotettix viridis* (Thomas) is an oligophagous grasshopper feeding on plants primarily in the host plant genera *Gutierrezia* and *Solidago* in Kansas. I used mitochondrial and microsatellite genetic markers to evaluate the diversification pattern and underlying evolutionary mechanisms of two putative host races of *H. viridis.* I also quantified host preferences, the degree of sexual isolation among putative host races and divergence in cuticular attributes to identify the nature and origins of initial barriers that isolated populations in the formative stages of divergence. mtDNA data revealed a star-shaped phylogeny, suggesting isolation in a single refugium ~110,000 years ago based on a molecular clock, followed by rapid population expansion. Microsatellite data reveal significant host-based genetic differentiation and structuring in *H. viridis* populations in Kansas, including a microsatellite locus under strong divergent selection. Neutral microsatellite loci did not reveal a differentiation pattern specific to host plant use. Significant host-based preferences by individuals that fed on two host plant groups were detected in host paired-feeding preference studies. No-choice mate selection experiments revealed preferences for individuals collected from the same host species independent of location with little mating observed between individuals from different host species. Significant differentiation in color and cuticular composition among different host plant races within the study area was also detected. Correlations between host choice, mate choice and phenotypic divergence were observed and this host associated divergence appears to have a

genetic basis. Based on the results of this study, I conclude that divergent selection for host plant use underlies observed sexual isolation among populations in this species. *Hesperotettix viridis* populations in Kansas that fed on *Solidago* and *Gutierrezia* species represent two incipient host races, early stages of diversification that could lead to speciation in insect herbivores.

HOST ASSOCIATED GENETIC DIVERGENCE AND SEXUAL ISOLATION IN THE GRASSHOPPER *HESPEROTETTIX VIRIDIS* (ORTHOPTERA : ACRIDIDAE)

by

TONY GRACE

M.S., Indian Agricultural Research Institute, New Delhi India, 2000

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

GENETICS
COLLEGE OF AGRIULTURE

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

Approved by:	Approved by:	Approved by:
Co-Major Professor **Dr. Susan J. Brown**	Co-Major Professor **Dr. Anthony Joern**	**C**o-Major Professor **Dr. Samantha M. Wisely**

ABSTRACT

Understanding evolutionary processes that structure genetic variation associated with lineage diversification and speciation is a central problem. Shifts in host plant use by insect herbivores and subsequent divergence in mating signals can lead to sexual isolation, especially when selection for specialization on different hosts restricts gene flow among populations. The grasshopper *Hesperotettix viridis* (Thomas) is an oligophagous grasshopper feeding on plants primarily in the host plant genera *Gutierrezia* and *Solidago* in Kansas. I used mitochondrial and microsatellite genetic markers to evaluate the diversification pattern and underlying evolutionary mechanisms of two putative host races of *H. viridis.* I also quantified host preferences, the degree of sexual isolation among putative host races and divergence in cuticular attributes to identify the nature and origins of initial barriers that isolated populations in the formative stages of divergence. mtDNA data revealed a star-shaped phylogeny, suggesting isolation in a single refugium ~110,000 years ago based on a molecular clock, followed by rapid population expansion. Microsatellite data reveal significant host-based genetic differentiation and structuring in *H. viridis* populations in Kansas, including a microsatellite locus under strong divergent selection. Neutral microsatellite loci did not reveal a differentiation pattern specific to host plant use. Significant host-based preferences by individuals that fed on two host plant groups were detected in host paired-feeding preference studies. No-choice mate selection experiments revealed preferences for individuals collected from the same host species independent of location with little mating observed between individuals from different host species. Significant differentiation in color and cuticular composition among different host plant races within the study area was also detected. Correlations between host choice, mate choice and phenotypic divergence were observed and this host associated divergence appears to have a

genetic basis. Based on the results of this study, I conclude that divergent selection for host plant use underlies observed sexual isolation among populations in this species. *Hesperotettix viridis* populations in Kansas that fed on *Solidago* and *Gutierrezia* species represent two incipient host races, early stages of diversification that could lead to speciation in insect herbivores.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude and adoration to my major profesors, Dr. Susan J. Brown, Dr. Anthony Joern and Dr. Samantha M. Wisely for their meticulous guidance incessant help, calm endurance and ceaseless encouragement during my tenure as a student under them. Not only are they excellent scientists from whom I learned a lot about research, they are good human beings from whom I learned a lot about life as well. Words will fall short in describing my heart felt gratitude to them for having hired me and for bestowing their trust in me. All of them have truly been excellent mentors. A special word of thanks to Sue Brown for the wonderful person she is and for everything she has done for me. There has never been a person who has impressed me more than her in my life.

I would also like to thank Dr. Richard W. Beeman for his suggestions in improving my work. I very much appreciate his trust in me and for having kindly agreed to serve on my committee.

I thank all members of the Brown, Joern and Wisely labs for their friendship and for having provided assistance in my research work. In particular, I would like to thank Jenny Apple, Barb, Michelle, Renata for the encouragement they provided. To my wife Ginny Antony for her loving support and assistance in difficult times without whom, I would never have come this far. Finally to my late mother and father who gave up everything they had and sacrificed their necessities to educate me and to ensure that I have a good future and oppurtunities that they never had.

Introduction

Understanding the origin and evolution of biodiversity and the underlying mechanisms of speciation are important issues in evolutionary biology. Whether from geographic isolation or adaptive divergence (ecological speciation) in sexual taxa, new species are formed when reproductive isolation occurs between two previously interbreeding groups. Some forms of speciation are somewhat controversial, mostly due to modes of speciation that are not well understood with respect to how reproductive barriers evolved to separate previously interbreeding groups (Diekmann and Doebell, 1999 ; Coyne and Orr, 2004). It is thus important to recognize opportunities for reproductive barriers that initially reduced gene flow between populations, the evolutionary and ecological factors responsible for creating divergence of lineages (Coyne and Orr 2004), and the mechanisms that underlie factors such as drift, selection and gene flow to understand speciation events. Recently, much effort has focused on the roles of selection and ecological processes to promote lineage diversification – ecological speciation – where studies of host plant use by insect herbivores provide critical case studies of underlying processes (Feder et al. 2003; Egan et al. *2008*; Nosil et al. 2008).

Most well known examples of speciation result from geographical isolation, leading to allopatric speciation. When populations are physically isolated from each other, genetic drift, natural selection and the lack of gene flow promote genetic divergence and ultimately, reproductive isolation of previously interbreeding lineages. However, sympatric populations are not so isolated geographically and may even coexist in the same locality, which reduces the likelihood that genetic differentiation that underlies initial stages of divergence will occur. A greater likelihood of gene flow between nearby populations fosters genetic mixing. Under

restrictive conditions, genetic divergence may still result under these situations, leading to sympatric speciation. Despite accumulating evidence, sympatric speciation is still sometimes a contentious issue, mostly because few examples exist. Sympatric speciation thus results from splitting of one evolutionary lineage into two without geographic isolation (Berlocher & Feder 2002). Some of the best examples of local divergence between lineages occur in phytophagous insects, where host race formation and eventual speciation has been reported often. Genetic divergence originates with populations shifting/dispersing to new or alternate host plants, where they evolve adaptive capabilities to successfully use new hosts. With limited gene flow and selection against hybrids because of the need to use different host plants, divergent selection of insect herbivores on different plants eventually leads to speciation (Bush, 1994; Larrson and Ekbom, 1995; Feder, 1998 ; Berlocher and Feder, 2002; Dre's and Mallet 2002). Because the process of speciation is ongoing, host races may retain characteristics responsible for reproductive isolation and thus provide a unique opportunity to study the speciation process (Itami et al. 1998). A third mode of speciation is the mixed mode where reinforcement occurs following secondary contact in sympatry after initial genetic divergence arises in allopatry (Marshall et. al. 2002; Smadja and Butlin, 2006; Noor, 1999; Servedio and Noor, 2003; Servedio, 2005; Coyne and Orr, 2004). This model assumes that initial genetic differentiation occurred in allopatry, possibly in different refugia, but the isolated populations come into contact before full reproductive isolation evolves. Mating proceeds, but local selection against hybrids may lead to the evolution of mechanisms promoting assortative mating, further reinforcing divergence.

Speciation and rates of divergence can be influenced by several factors, including genetic processes such as gene flow that acts to reduce genetic diversification and promote homogeneity

and genetic drift and selection that increase genetic divergence or heterogeneity. Spatial variation in selection could lead to local adaptations, the tempo of which depends on a balance between selection and the counteracting effects of gene flow (Slatkin, 1987**;** Nosil, 2005). Gene flow counteracts this divergence between populations with the exception of the situation of reinforcement where gene flow under limited conditions (is low enough to prevent homogeneity but high enough to promote reinforcement) could enhance selection to avoid hybridization (Nosil, 2005). Factors that affect gene flow among populations include dispersal ability, habitat persistence, habitat patchiness, population age etc. (Peterson and Denno, 1998).

It had been postulated as early as 1860 by Benjamin Walsh that several host specific herbivorous insects may have originated by shifting and adapting to new host plants. Since phytophagous host specific insects account for ~25-40% of all animal species, understanding the role of host shifts of speciation may be of tremendous value in understanding biodiversity in nature. Host races often represent an incipient stage of speciation for insect herbivores (Berlocher and Feder 2002). Well studied cases include the holly leaf miner *Phytomyza glabricola* (Scheffer and Hawthorne, 2007), apple maggot fly *Rhagoletis pomonella* (Feder et al. 1998, 2005), and goldenrod ball fly *Eurosta solidaginis* (Abrahamson & Weis, 1997). Current sympatric coexistence of two populations/species does not necessarily require that their specific origins are sympatric, or that barriers to gene flow separating two populations evolved under the same geographical locations or conditions (Michel et al. 2007). Lineages could have diverged allopatrically with the current coincident distributions representing secondary contact. Historical events (climatic events) easily could also separate populations, or alternately, facilitate local co-occurrence of previously separated populations. Pleistocene events fostered increased species diversification by separating populations through glacial barriers or as a result of geographical

shifts in favored habitat (Hewitt, 1996, 2000). Geographic variation currently observed could be a result of vicariance, leading to divergence in different glacial refugia (Runck and Cook, 2005). Biotic consequences of climatic changes once ameliorated resulted in substantial growth and expansion of several organisms resulting in them occupying newer habitats (Lessa, *et. al*., 2003) and thus expansion of species ranges. Under Pleistocene glaciation scenarios, phytophagous insect populations may have had to adapt to shifting host plant distributions or may have to shift to novel, locally co-occurring host plants if they could not track original host plant populations.

Hybrid zones

Hybridization in areas of secondary contact can be another important factor in understanding geographic variation. For species formed in allopatry, zones of secondary contact provide opportunities for hybridization, where reinforcing selection could lead to new species. The degree of genetic differentiation among divergent groups/host races/subpopulations can be examined in zones of secondary contact where lineages meet and produce hybrids (Barton and Hewitt 1985; Arnold 1997). Areas of secondary contact are thus maintained by a balance between dispersal and selection against hybrids (Nosil, 2005). Selection against hybrids can be caused by several factors. First, two lineages may be adapted to unique environments, resources or host plants (herbivorous insects) and hybrids may be less fit to adapt to the bimodal environments. Also alleles may be favored in their parental background and selection may act against hybrid genotypes via reinforcement (Barton and Gale, 1993). Populations that come into contact in a hybrid zone after developing in allopatry for some time could develop certain level of incompatibilities that get reinforced/diluted when they come together, depending on the extend to which discriminating mechanisms developed when they were separate. Some of these

mechanisms that might act to reinforce isolation could be selection against hybrids (low hybrid viability), lower fertility, higher mortality in case of interspecific mating or the development of mismatched mating characteristics (Servedio, 2005).

Reinforcement

Reinforcement is a mechanism for the evolution of reproductive isolation driven by natural selection that leads to speciation (Noor, 1995 ; Noor, 1999 ; Servedio, 2005). It is primarily the evolution and enhancement of existing discrimination mechanisms during mating to prevent interbreeding between populations in secondary contact. Because selection for reinforcement occurs in sympatry and not in allopatry, greater selection signatures in terms of mating preference or character divergence should be detected in areas where reinforcement occurs compared to allopatric areas. This has been demonstrated in several studies involving mate choice and hybrid viability assays including studies of walking sticks *Timema cristinae* and *Drosophila pseudoobscura* (Nosil et al. 2002; Noor, 1995). Still one of the biggest challenges for understanding the ultimate importance of reinforcement is the existence of relatively few well studied examples. Further study is required to unearth the factors that lead to reinforcement, and little is known about the biological or genetic factors that promote reinforcement except in a few cases. Ortiz Barrientos et al. (2004) demonstrated the genetic basis of reinforcement and concluded that reinforcement in their *Drosophila* system occurred as a result of the evolution of a different set of isolation/divergence mechanism than those in allopatric populations documenting that reinforcement is not just the enhancement of the existing barriers, and could be a result of evolution of new barriers.

Phenotypic differentiation

Some insects use color for identifying mates where as in others cuticular characteristics have been shown to aid mate recognition (Tregenza and Wedell, 1997) and mediate kin selection (Simmons, 1990). Variations in cuticular composition have been found significantly associated with levels of assortative mating in the grasshopper *Chorthippus parallelus* and in several other species and divergence in such mating signals could play a role in pre-mating reproductive isolation (Tregenza et al. 2000). In herbivorous insect species that feed on multiple plants, diet composition can influence the abundance and composition of hydrocarbons in the cuticle (Chapman et al., 2000). Cuticular chemicals are thus believed to play a role in communication during courtship, and increase behavioral isolation of host races. Thus, they may contribute to the process of speciation because divergence of such signals could lead to non-random mating between populations (Neems and Butlin, 1994). Variations in hydrocarbons have in fact been reported across a hybrid zone among sub-species of the meadow grasshopper *Chorthippus parallelus* (Neems and Butlin, 1994). Phenotypic differentiation could thus aid in mate selection or result as an after effect of mate selection and play a role in divergence.

Ecological divergence

Studies on herbivorous insects have resulted in better understanding of how ecological divergence can facilitate genetic diversification (Sword et al. 2005). Insect herbivores are classified based on their ability to utilize different plants. Specialist herbivores feed on single or few host plants (species or genera) and their life history and survivorship are tightly dependent on their host phenology and life history. Generalist feeding herbivores feed on a wide variety of host plants, often in different plant families, which allows them to use alternate hosts when the

favored host is not available (Mopper and Strauss, 1998). Considering their great diversity, it may not be possible to clearly define herbivorous insects as either true specialists or generalists (Bernays and Chapman 1994), but the distinction is generally useful. Based on their ability to feed on different hosts, a continuum of obligate specialists to facultative specialists and generalists exits (Orcutt et al. 2000). Even populations of the same species occupying diverse habitats vary in their host preference (Mopper and Strauss, 1998). Based on diet inventories from throughout species range, many insect species appear to be generalists (Fox and Morrow, 1981; Traxler and Joern, 1999). However, at finer scales, many populations show restricted diets that differ in host plants utilized from other populations (Fox and Morrow 1981; Mopper and Strauss, 1998; Traxler and Joern, 1999), a result expected for local adaptation to specific but different host plants (Traxler and Joern, 1999). Diet selection often varies among locations in response to ecological, physiological, behavioral and evolutionary forces. Constraints from physiological tradeoffs that limit performance on other hosts would reinforce the evolution of host specialization (Joshi and Thompson, 1995; Berenbaum and Zangerl, 1998). Positive correlations between host use and performance on existing host plants in a particular range could reinforce local adaptation to a particular host in that range and lead to host races associated with specific host plant species (Joshi and Thompson, 1995; Thomas and Singer, 1998).

Variations in host selection can affect mate choice through habitat and sexual isolation (Via, 1999; Nosil et al. 2006). Where divergent selection drives speciation, populations on different host plants should exhibit greater reproductive isolation than populations on the same host (Funk, 1998). Even though host associated differentiation is increasingly well documented and its role in ecological speciation is gaining wide acceptance, instances of host associated differentiation are relatively few compared to natural insect diversity (Stireman et al. 2005).

Most studies of host-based differentiation are reported for specialist herbivores such as in holly leaf miner *Phytomyza glabricola* (Scheffer and Hawthorne, 2007), apple maggot fly *Rhagoletis pomonella* (Feder et al. 2003, 2005), goldenrod ball fly *Eurosta solidaginis* (Abrahamson & Weis, 1997) and pea aphid *Acyrthosiphon pisum* (Via, 1999*).* Role of divergent selection and host adaptation as the major evolutionary force responsible for reproductive isolation and speciation in herbivorous insects have been on an increase as is evident from recent studies conducted on apple and hawthorn races of *Rhagoletis pomonella* (Feder *et. al*., 2003; Michel et al. 2007). Here, ecological divergence restricts gene flow among these two groups and in that of ecological divergence via natural selection in the sexual isolation of *Neochlamisus bebbianae* leaf beetles (Funk, 1998). Nosil (2007) also reported two walking stick host ecotypes of *Timmema cristinae* exhibiting strong selection in response to ecological divergence on different host plants. Unlike that observed in true specialists, processes that promote genetic divergence and differentiation are less well studied in oligophagous insect species. Thus my research on speciation in *Hesperotettix viridis* presented in the current study provides an exciting opportunity to study divergence in this context.

***Hesperotettix viridis*: background on the study species**

Grasshoppers are among the most polyphagous among insect herbivores (Joern 1979, Bernays and Chapman 1994, Traxler and Joern, 1999). While most grasshopper species are polyphagous or oligophagous only very few are monophagous (Otte and Joern, 1977, Traxler and Joern, 1999). *Hesperotettix viridis* (Thomas) is an oligophagous grasshopper native to North America (Joern, 1979, 1985, Traxler and Joern 1999) and feeds on composites in the family Asteraceae. Even though *H. viridis* is broadly distributed in North America, its distribution is

geographically patchy particularly in the east due to its dependence on a small number of host plants. Major hosts plants fed on by *H. viridis* typically include species from two genera, *Gutierrezia spp.* and *Solidago spp.* Currently *H. viridis* is more commonly encountered in the western part of its distribution. Individuals spend most of their life activities such as feeding, searching for mates and copulating, thermoregulation and avoiding predators on the host plant (Parker, 1983 ; Pfadt, 1994 ; Traxler and Joern, 1999). Adults have functional wings but are only weak fliers. *H. viridis* shows low vagility and even short distances between hosts plants could serve as a barrier to dispersal (Parker 1983)**.** This grasshopper species has a single generation/year and overwinters as eggs deposited in soil (Pfadt, 1994). Eggs are laid as pods in soil beneath the host plant at a depth of one inch (Traxler and Joern, 1999). There are 5 nymphal instars and all stages feed on the same host plant. Nymphs emerge in late May to early June and adults emerge in mid to late July (Traxler and Joern, 1999). Even though *H. viridis* has been reported to feed on ~35 different host plants throughout its distributional range (Pfadt 1994), local populations typically exhibit restricted diets of 1-2 host species and populations often differ with location in host plants utilized (Traxler and Joern, 1999). Such patterns are consistent with an interpretation of local adaptation (Mopper and Strauss, 1998). Current feeding patterns could be a result of recent or ancient host shifts via a single event/split or as a result of multiple events occurring independently across its distribution.

Reciprocal transplant experiments between alternate host species conducted with *H. viridis* in the field (Traxler and Joern, 1999) yield results consistent with an interpretation of local adaptation to unknown general ecological differences (Mopper and Strauss, 1998) when populations on same hosts but different locations were compared. But populations on different host plants even at the same site support a hypothesis of host race formation possibly leading to

lineage diversification. Reciprocal transplant experiments showed that individuals collected from *Solidago mollis* performed better on *S. mollis* than on *Gutierrezia sarothrae* and vice versa, suggesting the existence of adaptation to the respective host plant despite the fact that two host plants occur together at the same site. Preconditioning on the opposite host did not alter results, further suggesting that local processes were important in reinforcing differentiation in host plant use. The fact that grasshoppers from *G. sarothrae* exhibited more difficulty in using *S. mollis* (alternate host) than individuals collected from *S. mollis* indicated a more constraining or costly mechanism associated with host specialization on *G. sarothrae* (Traxler and Joern 1999). An alternate interpretation is the possibility of *G. sarothrae* is the historical primary host followed by a possible recent host shift to *Solidago* spp. No underlying mechanism involved in any possible shift has been identified to indicate that the individuals collected from *G. sarothrae* are not able to utilize the new host plant as efficiently as individuals collected from *S. mollis*. If the host shift hypothesis is correct, individuals collected from *S. mollis* from a possible recent shift from *G. sarothrae* still have the capability to feed reasonably well on both host plants. Studies conducted by Traxler and Joern (1999) point to local factors being more important, suggesting a scenario of ecological speciation.

Using AFLP markers, Sword et al. (2005) examined genetic relationships among five populations of *H. viridis* that fed on *G. sarothrae* and *S. mollis*, and found significant genetic differentiation among populations feeding on two host plants even when present in the same locality. Genetic analysis using AFLP's suggested that differentiation occurred much earlier in the evolutionary history of *H. viridis*. Recently, Apple et al. (submitted) using AFLP markers extended these results and found significant host-based differentiation among 23 populations from 7 states of *H. viridis* on 3 different host plants. Significant differentiation was observed at

both selected and neutral loci suggesting a significant role for selection and genetic drift in the genetic differentiation observed. Significant differentiation was observed among *Gutierrezia* and *Ericameria* feeding forms vs. *Solidago* feeding forms where as no significant differentiation was observed among *Gutierrezia.* vs. *Ericameria*.

Primary goal of study

From these studies, it appears that selection and to a limited extent genetic drift may have both played a role in structuring observed variation among *H. viridis* host variants at a broad geographic scale. However, fine-scaled genetic studies within a well detailed area across zones of overlap are lacking. Assessing variation across zones of overlap will be valuable in testing boundaries of speciation and to ascertain the extent of reproductive barriers among the putative host races. No studies have targeted feeding preference, reproductive isolation, phenotypic differentiation or gene flow among the host variants across a zone of overlap, each of which provide important evidence in addressing speciation mechanisms and in understanding evolutionary forces that drive such diversification. Some basic questions include: Do correlations exist between host choice and mate choice? Is there evidence for phenotypic differentiation in association with host use, and does it contribute to sexual isolation? Is there genetic structure associated with host choice and mate choice? Such questions are addressed in this study.

After extensive collecting throughout the U.S., we identified a zone of syntopy in Kansas where host plants, *Gutierrezia spp.* and *Solidago spp.* used by *H. viridis* coexist. The distribution of *H. viridis* in Kansas is such that eastern distributions only encounter *Solidago* populations, western Kansas populations feed on only *Gutierrezia*, and the central zone of syntopy includes populations of both host plant populations coexisting together. Syntopic populations collected from both host plants in the central region of Kansas and populations collected on *Gutierrezia sp.*

on the western portion of the study transect are well within the geographic range of the major host plant *Gutierrezia spp.,* whereas the populations collected from *Solidago spp.* on the eastern portion are outside the distribution of *Gutierrezia spp*. This distribution of host plants and grasshoppers provides an excellent opportunity to examine host-based diversification in the context of different selection pressures associated with feeding on one host versus when both host plants coexist inside and outside the distribution of the major host plant. The lack of the *Gutierrezia* in eastern Kansas could potentially reflect situations that occur when organisms get isolated in refugia with a related but different host plant, requiring shifts and adaptation to new nutritional selective pressures. Even though no obvious physical barrier to gene flow exists on the eastern side of the distribution, the total lack of the alternate host could serve as an ecological barrier and shifts by *H.* viridis onto this host species could promote divergence. Thus, the presence/absence of the major host plant could either accelerate or decelerate the tempo and mode of diversification; this Kansas study system provides a good backdrop to test various models of speciation.

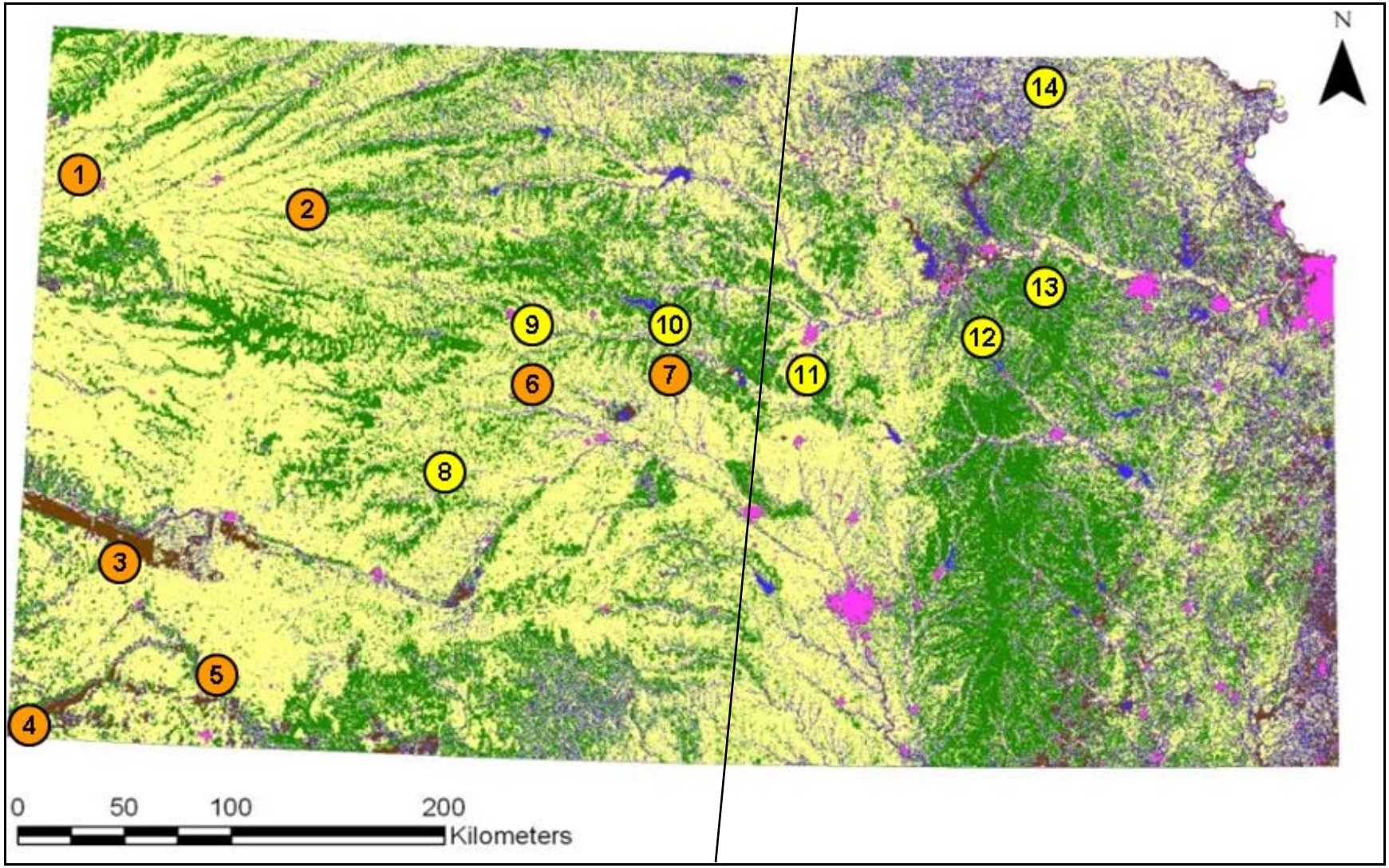
Genetic differentiation and gene flow: I will analyze host plant based genetic differentiation in the grasshopper *Hesperotettix viridis* using microsatellite markers and mtDNA. This will determine if divergent selection contributes to genetic differentiation and compare and contrast variations and gene flow patterns at both neutral and selected loci to separate the effects of selection and drift in differentiation of *H. viridis* populations. The zones of syntopy provide opportunities to assess genetic differences in *H. viridis* populations, thus providing an opportunity to examine mechanisms shaping species boundaries in the zone of overlap. The distribution of host plants provide exciting possibilities for studying gene flow between

populations of *Hesperotettix viridis* on alternate hosts in allopatry as well as that of co-existing populations in the zone of syntopy. I will ascertain whether sympatric co-existence observed in the middle of the distribution in Kansas are: (a) populations that have come into secondary contact after initial divergence elsewhere, or (b) if subgroups of *H. viridis* on different hosts are the result of a recent host shifts in the zone of syntopy, suggesting that sympatric speciation is in progress. Possibilities also include populations expanding from *Gutierrezia* and *Solidago* systems on the western and eastern sides to form a *Gutierrezia/ Solidago* zone of syntopy (a stable hybrid zone in the center of the transition between the distribution of host species), or a continuum or progressive expansion of populations in the study area involving an expansion from the major host plant *Gutierrezia to the Gutierrezia/Solidago* system in the center and to alternate host *Solidago* in the east. Sympatric speciation in the zone of syntopy followed by outward expansion of populations to the periphery is also another possibility.

Feeding preferences, sexual isolation and phenotypic differentiation: I intend to test the relative contributions of host plant preferences and phenotypic differentiation in the sexual isolation of *H. viridis* host populations in Kansas. Feeding preference trials would assess the extent of host association among different populations and assess if character displacement occurs in the zone of syntopy. Character displacement occurs when there is higher feeding preference divergence among the two host plant groups in syntopy than in locations where the distributions do not overlap. This would test for significant preference for the home/native host which implies that host plant use drives ecological divergence. In such cases, if adaptation to specific host plants drive sexual isolation, higher magnitude of sexual isolation would occur among ecologically different pairs of populations that among ecologically similar populations. Here I test for strength

of mating isolation among ecological divergent and similar hosts. If mating trials indicate host based assortative mating, this would suggest that ecological divergence is also reinforced by pre-mating isolation. Since variations in cuticular composition have been found significantly associated with levels of assortative mating in several insects, in order to investigate possible phenotypic differentiation associated with ecological divergence, I will also characterize variations in cuticular characteristics and color morphology among host races of *H. viridis* using near-infra red spectroscopy. This study would address if phenotypic divergence has any role in promoting divergence of *H. viridis* populations.

Thus I propose a fine scale study within Kansas across a zone of syntopy of *H. viridis* to assess the host plant association among populations (feeding preference), the extent of sexual isolation (mate choice) and phenotypic differentiation among populations on alternate hosts and a genetic study using microsatellite markers and mtDNA to further clarify differentiation patterns and to understand the evolutionary forces involved in diversification of *H. viridis*.



1:Sherman G ; 2: Hoxie-G ; 3: Hamilton-G ; 4: Cimmaron-G ; 5: Seward-G ; 6: Ellis-G ; 7: Russell-G
8: Ness-S ; 9: Ellis-S ; 10: Russell-S ; 11: Ellsworth-S ; 12: Marrion-S ; 13: Konza-S ; 14: Marysville-S

Fig 1. Sampling locations for *Hesperotettix viridis* within Kansas. 1-7 *Gutierrezia* populations; 8-14 *Solidago* Populations. Populations 1-10 are within the distribution of major host plant *Gutierrezia sp.* ; Populations 6.7,9,10 are in zone of syntopy.

Chapter II

Host-Associated Genetic Divergence in the Grasshopper *Hesperotettix viridis* (Orthoptera: Acrididae) - A Population Genetic study using Microsatellites and mtDNA

Running head: Genetic diveregence in *Hesperotettix viridis*

Tony Grace, Anthony Joern, Susan J. Brown and Samantha M. Wisely*
Division of Biology, Kansas State University, Ackert Hall, Manhattan, KS 66506

Keywords: ***Keywords:*** *Hesperotettix viridis*, snake weed grasshopper, mitochondrial DNA, microsatellite markers, local adaptation, host associated divergence, biogeography, divergent selection, speciation.

*Corresponding author: Samantha M. Wisely, Division of Biology, Kansas State University, Ackert Hall, Manhattan, KS 66506 Phone (785) 532-6615 ; Email: wisely@ksu.edu

This chapter is being prepared for publication submission.

Abstract

Understanding the evolutionary processes that structure genetic variation is important when studying speciation mechanisms. Variations in diet choices can often translate into assortative mate choice from selection for habitat preferences and lead to genetic differentiation. Studies involving herbivorous insects and host plant associations have provided important insights into how ecological divergence can facilitate genetic diversification in numerous taxa. In such cases, selection in contrasting directions restricts gene flow among populations and drives previously interbreeding populations to ecological speciation. Mitochondrial and microsatellite markers were used to evaluate the contributions of various evolutionary forces in diversification pattern of two putative host races of the grasshopper, *Hesperotettix viridis* in Kansas. *H. viridis* is an oligophagous species feeding on plants primarily in the host plant genera *Gutierrezia* and *Solidago* in Kansas. Mitochondrial DNA data revealed a star shaped phylogeny suggesting isolation in a single refugium followed by rapid population expansion. A recent divergence time estimate according to the molecular clock of ~110,000 years corresponds to the Sangamon interglacial period. Microsatellite data reveal significant host based genetic differentiation and structuring in *H. viridis* populations in Kansas. We identified a microsatellite locus under strong divergent selection that appeared as an outlier in multiple population comparisons. Neutral microsatellite loci did not reveal a differentiation pattern that correlated with host species. The results of the study support the role of rapid divergent selection as the evolutionary force behind the observed host-based differentiation pattern.

Introduction

The central problem in speciation is to understand the evolutionary forces that initially reduced genetic exchange between populations (Coyne and Orr 2004). Speciation can occur as a result of adaptation to contrasting environments or habitats, or occur without adaptation (Butlin et al. 2009). Recently much effort has focused on the roles of selection and ecological processes in lineage diversification (ecological speciation) where studies of host plant use by insect herbivores provide critical case studies of underlying processes (Feder et al. 2003; Egan et al. *2008*; Nosil et al. 2008). The close association between insect herbivorous and their host makes them vulnerable to disruptive selection following divergence in ecological traits (Mopper and Strauss, 1998). It had been postulated as early as 1860 by Benajamin Walsh that several host specific herbivorous insects may have originated by shifting and adapting to new host plants. Since phytophagous host specific insects account for ~25-40% of all animal species, understanding the role of host shifts in speciation is of tremendous value in understanding biodiversity in nature.

Host races represent a stage in the speciation process where adaptation to different environments (Scheffer & Hawthorne, 2007) promotes reproductive isolation among previously interbreeding populations. Under sympatric speciation scenarios, host shifts result when directional selection drives contiguous populations towards disruptive selection (selection against hybrids) thereby promoting adaptive divergence (Doebeli, 1996; Geritz *et. al*., 1999; Via, 1999; Doebeli and Dieckmann, 2000,). Populations shifting/dispersing to new or alternate host plants must evolve adaptive capabilities to successfully use new hosts. With limited gene flow and selection against hybrids, divergent selection could eventually lead to genetic differentiation and speciation (Bush, 1994 ; Larrson and Ekbom, 1995 ; Feder, 1998 ; Berlocher and Feder,

2002 ; Dres and Mallet, 2002). Because the process of speciation is ongoing, host races may retain characteristics responsible for reproductive isolation and thus provide a unique opportunity to study the speciation process (Itami, et al. 1998). Divergence in such cases could occur due to recent or ancient host shifts that occurred within a single refugium or in different refugia as a single event/split or as a result of multiple events across the distribution of a species. Thus current sympatric coexistence of host races/species does not necessarily require that their specific origins are sympatric, or that barriers to gene flow separating two populations evolved under the same geographical locations or conditions (Michel, *et. al*., 2007). Lineages could have diverged allopatrically with the current coincident distributions representing secondary contact. For species formed in allopatry, such zones of contact provide opportunities for hybridization and testing grounds for species boundaries. Alternatively population divergence could initially occur in allopatry followed by reinforcement in sympatry. Thus both mechanisms could act to affect one speciation event (Feder *et. al*., 2003 ; Michel et al. 2007).

The importance of ecological contributions to divergence and speciation has gained recognition in recent years because host shifts have increased diversification of species through divergent selection or genetic drift. Ecological divergence and adaptation as a result of diet variation plays an important role in promoting reproductive isolation and speciation and positive associations exist between the two (Funk, et al. 2006). Variations in diet choices can often translate into mate choice from selection for habitat preferences and lead to genetic differentiation. Thus, studies on herbivorous insects can provide excellent insights into how ecological divergence can facilitate genetic diversification (Sword et al. 2005). Based on diet inventories from throughout a species range, many insect species appear to be generalists (Fox and Morrow, 1981; Traxler and Joern, 1999). However, at finer scales, many populations show

restricted diets that differ in host plants utilized from other populations (Fox and Morrow, 1981 ; Mopper and Strauss, 1998 ; Traxler and Joern, 1999), an expected result for local adaptation to specific but different host plants (Traxler and Joern, 1999). Constraints from tradeoffs could limit performance on other hosts and thus would reinforce the evolution of host specialization (Joshi and Thompson, 1995; Berenbaum and Zangerl, 1998) after host shifts and result in divergence. However, most studies of host-based differentiation are reported for specialist herbivores such as in holly leaf miner *Phytomyza glabricola* (Scheffer and Hawthorne, 2007), apple maggot fly *Rhagoletis pomonella* (Feder et al. 2003, 2005, 2007), and goldenrod ball fly *Eurosta solidagoidaginis* (Abrahamson & Weis, 1997). Less well understood is the role of these mechanisms in the ecological speciation of generalist herbivores.

In the current study I analyzed host plant based genetic differentiation in the grasshopper *Hesperotettix viridis* using microsatellite markers and mtDNA. Grasshoppers are among the most polyphagous insect herbivores (Joern 1979; Bernays and Chapman 1994; Traxler and Joern, 1999). While some grasshopper species are oligophagous only very few species are truly monophagous across their geographic range (Otte and Joern, 1977; Traxler and Joern, 1999). *Hesperotettix viridis* is oligophagous, feeds on composites (Asteraceae), is a native of North America (Joern, 1979, 1985, Traxler and Joern 1999) and broadly distributed. Even though *H. viridis* has been reported to feed on ~35 different host plants throughout its distributional range (Pfadt 1994), local populations typically restrict diets to 1-2 host species (Traxler and Joern, 1999), and populations often differ in host plants eaten. Major hosts eaten by *H. viridis* typically include species from three genera : *Gutierrezia*, *Solidago* and *Ericameria.* Reciprocal transplant experiments conducted on *H. viridis* (Traxler and Joern, 1999) suggest local adaptation (Mopper and Strauss, 1998) even within putative host races. But populations on different host plants even

at the same site support a hypothesis of host race formation and lineage diversification (Sword et al. 2005). Unlike that observed in true specialists, processes that promote genetic divergence are less well studied in oligophagous insect species. *H. viridis* thus provides an exciting opportunity to study genetic divergence in this context and for unraveling the factors that drive diversification.

Several ecological and non-ecological forces structure genetic variation in populations. Mutations, random genetic drift (Kimura 1996), selection (Funk, 1998; Egan et al. 2008) and population bottlenecks/ founder events (Coyne and Orr, 2004) all promote differentiation. Climatic changes during the Pleistocene could have isolated *H. viridis* populations in different refugia with different host plants and promoted differentiation via genetic drift, or host shifts could have occurred within the same refugia via divergent selection. Genetic drift could have structured divergence among populations of *H. viridis* as previously observed for the grasshopper *Melanoplus oregonensis* (Knowles and Richards 2005). Thus understanding the evolutionary processes that structure variation is important when studying speciation mechanisms and host plant based differentiation in *Hesperotettix viridis*.

Previous studies of ecological divergence have found restricted gene flow and natural selection in the sexual isolation of *Neochlamisus bebbianae* leaf beetles (Egan et al. 2008), host ecotypes of walking stick *Timmema cristinae* (Nosil et al. 2008), leaf miner *Phytomyza glabricola* (Scheffer and Hawthorne, 2007) and apple maggot fly *Rhagoletis pomonella* (Feder et al. 2003). Gene flow could thus promotes divergence (low gene flow) or alternatively constrain divergence (high gene flow). Estimating gene flow patterns among diverging populations, particularly across zones of overlap, will be helpful in clarifying the taxonomic status of the host variants and prove valuable in assessing the strength of reproductive barriers among diverging

lineages. Apple et al. (submitted) found significant host associated differentiation among 23 populations of *H. viridis* from 7 western states in the U.S. on three different host plants. Significant differentiation was observed at both selected and neutral loci, and a role for selection and genetic drift in the differentiation pattern was inferred. From this, it appears that selection and genetic drift have both played a role in structuring variation among *H. viridis* host variants at a broader scale. However detailed and fine scaled genetic studies within a well detailed area across zones of overlap are lacking. Assessing genetic variation and patterns of gene flow across zones of overlap will be valuable in testing boundaries of speciation in *H. viridis*

After extensive sampling, I identified a zone of syntopy in Kansas where two distinct host forms that feed on host plants in the genera *Gutierrezia* and *Solidago* coexist. We conducted a fine-scale population genetics analysis to characterize the genetic structure and gene flow patterns among *H. viridis* populations across Kansas, where a zone of syntopy with grasshopper populations on both hosts (*Solidago spp.* and *Gutierrezia sarothrae*) co-exist, spanned geographically on both sides of the distribution by populations on single but different host plants. The distribution of *H. viridis* populations and their host plants in Kansas (West: *Gutierrezia;* syntopy: *Gutierrezia* & *Solidago*; East: *Solidago*) provide an excellent opportunity to look at host-based diversification in the context of different selection pressures associated at sites with feeding on one host versus sites when host plants coexist, as well as within and outside the distribution of the major host plant *Gutierrezia*. The lack of the major host, *Gutierrezia*, in eastern Kansas could potentially reflect situations that occur when organisms get isolated in refugia with a different host plant, requiring shifts and adaptation to new nutritional selective pressures. Even though no obvious physical barrier to gene flow exists along the eastern side of the transect lack of the alternate host would serve as an ecological barrier and promote genetic

drift or divergent selection. The presence/absence of the major host plant could either accelerate or decelerate the tempo and mode of diversification and thus our study area provides a good backdrop to test different modes of diversification.

I will determine if divergent selection contributes to genetic differentiation and compare and contrast variations and gene flow patterns at both neutral and selected loci to separate the effects of selection and drift in differentiation of *H. viridis* populations. Several possibilities could occur in the zone of syntopy in Kansas and will be addressed. Using mitochondrial and microsatellite markers, the following hypotheses will be addressed: (a) these populations are not genetically different even though they appear different phenotypically; this area is a zone with genetic exchange between two "phenotypes" of the same species (panmixia). (b.1) The study area is a stable zone of contact between the two taxa where a low level of gene flow occurs between the two genetically diverged populations. This would suggest that these are two host races with limited genetic exchange between them, exhibiting incipient speciation. (b.2) the two phenotypes represent good species and are genetically separated from each other. Although taxa coexist at the same site, no genetic exchange occurs. (b3.) Sympatric host race formation in the zone of syntopy in Kansas followed by population expansion to the periphery. For the last three possibilities, we will test if populations may have diverged by genetic drift or selection in allopatry in two refugia followed by secondary contact or a host shift in a single refugium or location followed by rapid population expansion and divergent selection.

Materials and Methods

Insect Sampling.

Hesperotettix viridis individuals were collected from populations along four transects in Kansas that crossed the distributional limits of possible host plant species. In the zone of syntopy, two replicated populations of *H. viridis* feeding on multiple host plants (*Gutierrezia sp.* and *Solidago spp.*) in the same location were collected. In addition, we also collected *H. viridis* individuals from populations encountering only one of their major host plants, *G. sarothrae* (5 populations) and *Solidago spp.* (5 populations) on both western and eastern sides of the two syntopic zones respectively. Populations collected from both host plants in the zone of syntopy and populations collected on *Gutierrezia sp.* on the western side are well within the geographic range of the major host plant *Gutierrezia sp.,* whereas the populations collected from *Solidago sp.* on the eastern side were outside the distribution of *Gutierrezia sp*. ~30 individuals from each of 14 populations were collected.

After observing and recording activity of each insect collected on its host plant (feeding, mating, resting on host plant), each individual was collected and placed separately in vials containing 95% ethanol and then stored at -20^0 C until further use. When sampling, *H. viridis* individuals were collected from their respective host plants by hand or using sweep nets. Dried plant voucher specimens as well as wet specimens stored in alcohol for gut content analysis to ascertain host association were collected at every site and archived for further studies later.

***DNA extraction and marker optimization* :** DNA was extracted from muscle tissue collected from the femur of the hind leg using DNeasy Blood and Tissue Kit from QIAGEN Inc.

Concentrations of DNA extracts were measured using a nanodrop spectrophotometer and was stored at -20C till further use.

Co1 Mitochondrial DNA Amplification and Sequencing: Occurrence of mitochondrial pseudogenes in nuclear genomes of grasshoppers is a common phenomenon (Zhang & Hewitt, 1996; Sword et al. 2007) and has been reported in several acridid subfamilies (Bensasson et al. 2000; Sword et al. 2007). To avoid such a possibility for the mtDNA part of the study, we tested 4 different gene regions of mtDNA and characterized one 373 bp region of the *Cytochrome oxidase* gene devoid of intra-individual heterogeneity for use in phylogenetic analysis. We also tested for the presence of *Wolbachia* to eliminate the possibility of a mitochondrial selective sweep by using 2 sets of *Wolbachia* specific gene primers; we found no signs of *Wolbachia* infection (Grace et al., 2009, unpublished data). We amplified 373 bp of the *Co1* gene from 3-5 individuals in 13 populations using primers CO1aHVF (GAG CAC CGG ATA TAG CAT TTC CAC GA) and CO1aHVR (AAT AGG ATC ACC TCC TCC TGC AGG AT). Amplifications were carried out in an Eppendorf Master Cycler (Eppendorf Inc) in 25 µL volume. The reaction mixture contained 12.5 µL of water, 1x Taq buffer, 1.5 mM MgCl2, 0.2 mM dNTPs, 0.25 µM of each primer, 1 U Taq DNA polymerase and 50–100 ng DNA. Cycling conditions were: denaturation at 94 °C for 3 min followed by 35 cycles of 95°C for 30 s, 60°C for 45 s and 72°C for 30 s, and a final extension at 72°C for 15 min. Amplicons were separated on a 1% agarose gel, purified using PCR purification columns (Qiagen), and sequenced in both directions on an ABI 3730xl automated sequencer at the Advanced Genetic Technologies Center (AGTC) at the University of Kentucky. Sequences were corrected and aligned using SEQUENCHER 4.1.2 (Gene Codes Inc).

Microsatellites: We developed 7 polymorphic microsatellite markers from genomic DNA (Grace et al. 2009) enriched for CA and mixed (GA, GT, GTT etc) microsatellite repeats as described in Hamilton et al. (1999). Primer sequences and PCR conditions were as described in Grace et al. (2009). Individuals from 14 populations (n = 360) were amplified. The PCR fragments that were fluorescently labeled (HEX or FAM, Applied Biosystems) in a 3 primer PCR reaction (Schuelke, 2000) were visualized on an ABI 3730 DNA analyzer (Applied Bio-systems, Foster City, CA) using a ROX-400 size standard (Applied Bio-systems). Alleles were scored using GENEMARKER ver. 1.71 software (Soft Genetics Inc).

Data analysis

MtDNA Analysis: Diversity estimates and time of divergence estimates

Haplotype networks better depict relationships among populations within a species than phylogenetic trees (Trontelj et al., 2005; Bandelt *et .al.,* 1999) when genetic distances among individuals are small and in cases where genetic exchange among groups could result in a reticulate pattern (Trontelj et al. 2005). A median joining network was constructed to depict relationships among haplotypes using the program Network *ver*. 4.501 (http://www.fluxus-engineering.com/sharenet.htm). Character states were assigned equal weights, and the homoplasy factor, ε, was given a value of 0, which is the most parsimonious value for the parameter.

Arlequin ver. 3.1. (Excoffier et al. 2005) was used to calculate AMOVA, haplotype diversity (h), nucleotide diversity (π), Fu's F_s and Tajima's D. We used haplotype and nucleotide diversity estimates to interpret signatures of recent demographic events such as range expansion. Tajima's D and Fu's F_s were used to detect signatures of past demographic events on

current population stability or growth and expansion. Genetic variance among groups of populations at geographic and host plant level was partitioned using AMOVA (F_{CT} values). To test for geographic partitioning, populations were split into 3 regions: Eastern, Central and Western Kansas. To test for partitioning of variance based on host plant association, grasshopper samples were split in to two groups based on which host plant the sample was collected from *(Gutierrezia and Solidago*). AMOVA among these groups and diversity indices within each group were calculated separately to infer variability at different levels of organization.

Divergence times among haplotypes obtained in Network ver 4.501 were estimated using the Brower arthropod sequence divergence rate of 2.3% per million years for *Co1* gene (Brower, 1994), which has also been used in several related species including *Melanoplus sp*. (Carstens and Knowles, 2006). Calculations were based on the rate of 1 mutation/substitution every 116,877 years for the 373 bp of the CO1 sequence.

Microsatellite Analysis

Statistical tests were conducted to test for deviations from Hardy- Weinberg equilibrium and to check for linkage disequilibrium using ARLEQUIN ver. 3.1 (Excoffier et al. 2005). Pairwise F_{ST} values (Wright, 1951; Weir and Cockerham, 1984) were estimated using ARLEQUIN ver. 3.1 (Excoffier et al. 2005) and significance was estimated by permutation tests (10,000 randomizations). Private alleles in the two host plant groups were estimated using GDA ver. 1.1 (Lewis and Zaykin, 1999). Allelic richness in each population and inbreeding coefficient (F_{IS}) values were estimated using F_{STAT} ver. 2.9.3.2. (Goudet, 2001).

A coalescence-based simulation approach in DETSEL ver 1.0 (Vitalis et al. 2003) was used to identify any outlier loci that were potentially affected by selection. The expected joint

distributions of populations were generated by performing 100,000 coalescent simulations for each pair wise comparison of populations. The default nuisance parameters in the program were used to generate null distributions with similar number of allelic states as in the observed data set. Loci that fell outside the 0.95% probability region compared to the simulated data points were considered as outliers and potentially affected by selection.

Migration rates (*M*) and effective population sizes **(*Ne*)** were estimated using the coalescent approach as implemented in MIGRATE (Beerli & Felsenstein, 2001). Markov chain Monte Carlo-based (MCMC) maximum likelihood procedure was used for the estimation of population genetic parameters where ten short and three long chains were run with a short sampling increment of 20. In the burn in phase, the first 10,000 trees were discarded for each run. The analysis was continued until convergence of parameters was observed.

Structure *ver*. 2.2 (Pritchard et al. 2000) was used to assign individuals from all 14 populations to a predetermined number of clusters (K= 1-6) based on multilocus microsatellite data. A burn-in period of 25,000 steps followed by 50,000 iterations under the admixture model was undertaken for each run. We found that a burn-in length of 25,000 and MCMC of 50,000 was sufficient and longer burn-in or MCMC did not change the results significantly. 10 replicates for each run were performed. Log-likelihood scores were averaged across the 10 runs and compared to determine the posterior probability (LnPD) of each K. The highest Ln(PD) value or the value at which the LnPD did not increase any further and tapered off was selected as the K value corresponding to the number of clusters present in the data. Individual assignment tests for all populations were also conducted in GENALEX 6.0 (Peakall and Smouse, 2006)

where individuals were either assigned to *Gutierrezia* cluster or to a *Solidago* cluster. Analysis of molecular variance (AMOVA) undertaken using Arlequin ver 3.1 was used to partition genetic variation among different levels of among host plant among populations and within populations. Separate analyses were also done for host plants within the zone of syntopy and host plant populations outside the zone of syntopy. The correlation of Nei's genetic distance values and the natural log of geographic distance were performed for microsatellite data with a Mantel test. Probabilities were estimated in IBD 1.52 (Bohanak, 2002) and significance was tested using 10,000 randomizations of the data.

We used the Wilcoxon sign test in Bottleneck 1.2 (Piry et al. 1999) to test for signs of bottlenecks/ founder events in the recent history of the populations. This program estimates for each population the heterozygosity expected from the observed number of alleles that is then compared to the observed heterozygosity to ascertain whether there is a heterozygosity excess or deficit. This is based on the principle that populations that have experienced a recent reduction of their effective population size would exhibit a corresponding reduction of allelic diversity and heterozygosity. In this case, the allelic diversity is reduced faster than the heterozygosity, i.e. the observed heterozygosity is larger than the heterozygosity expected from the observed allele number. (Piry et al. 1999)

We used the Seqboot option for gene frequencies in PHYLIP ver 3.68 (Felsenstein, 2005), to generate 1000 bootstrapped matrices of Nei's genetic distances. Using the program NEIGHBOR, 50% majority rule neighbor-joining trees were constructed and a consensus tree was selected using the program CONSENSUS in PHYLIP 3.68 (Felsenstein, 2005). Neighbor-joining trees constructed with all microsatellite loci were compared to trees based on neutral loci and the selected locus.

Results

MtDNA analysis

373 bp sequence of the *Co1* gene were used for all analyses. 30 sites were polymorphic and informative and 29 different haplotypes were observed across 13 populations. Median joining analysis revealed a star shaped haplotype network where the most common haplotype was surrounded by less frequent haplotypes (Fig 1). 22 of 55 individuals shared the same haplotype which spanned across the two host plant groups (12 in *Gutierrezia sp.* and 17 in *Solidago sp.).* Out of the 29 total haplotypes, only 2 haplotypes were common across the host plant groups and 27 haplotypes were rare and not shared among host plant groups. Divergence time estimates from the most common haplotype (considered as the ancestral node) to other haplotypes (descendant nodes) indicated a divergence time 110,502 ± 26,199 years ago (Fig 1).

Diversity estimates were analyzed for individuals collected from each of the 2 host plant groups. Both groups showed high haplotype diversity (h = 1.0) and low nucleotide diversity (π *in Gutierrezia* = 0.004 and *Solidago* = 0.005*)*. Nucleotide and haplotype diversity levels were similar and not significantly different across host plant groups. Tajima's *D* values were negative for both groups (*D*; *Gutierrezia* = -2.124 and *Solidago* = -1.961) as were the Fu's *F* values (*F*; *Gutierrezia* = -27.76 and *Solidago* = -26.92) (Table 1). These patterns of diversity are indicative of recent population growth and expansion. Partitioning of molecular variation among the two host plant groups (AMOVA) revealed no significant variation among the two groups with only 1.48% of the total variation being attributed to among host plant groups where as 10.26% of the variation was among populations within groups and 88.26% of the variation within populations (Table 1). Similarly, comparisons among three other grouping in the study zone; a) *Gutierrezia*

vs. *Solidago* populations in the zone of syntopy b) zone of syntopy *Gutierrezia* vs. other *Gutierrezia* populations c) zone of syntopy *Solidago* vs other *Solidago*) revealed no significant variation among the groups compared (Table 1). Results of partitioning variance by geographic locations (central zone with both *Solidago* and *Gutierrezia*, western zone *Gutierrezia* and eastern zone *Solidago*) were similar to host plant-based differentiation (data not presented).

Microsatellites

Selection

Because previous studies point to host associated divergence and adaptation (Traxler and Joern, 1999; Sword et al. 2005), we tested if loci were under divergent selection. Loci linked to genes affecting fitness will exhibit higher level of differentiation compared to neutral regions. One locus, Hvir97, was under strong selection and appeared as an outlier in all pair wise comparisons of different host population differentiation. Locus Hvir97 fell outside the 95% confidence interval of genetic differentiation derived through simulation under a neutral mode of divergence (Fig 2).

Microsatellite Diversity, Hardy-Weinberg Equilibrium and Linkage Disequilibrium

All seven loci were polymorphic. The number of individuals in each population varied from 8-47 and the number of alleles varied from 5-52 across all loci (Table 2). All loci except the locus under selection showed similar levels of allelic diversity. There was a drastic reduction in allelic numbers at the selected locus. Average observed and expected heterozygosities were 0.57 to 0.88 across populations and loci. Significant deviations from Hardy Weinberg equilibrium (HWE) was observed in 68 of the 105 tests of comparisons in populations and all loci. Significant heterozygote deficiencies were observed across populations which resulted in the

deviations from Hardy-Weinberg equilibrium, and global tests for Hardy Weinberg Equilibrium showed significant deviations at all loci. Exact test of linkage disequilibrium (LD) using Fishers exact probability test revealed linkage disequilibrium in 134 of 316 tests conducted for pairs of loci in population level tests. No consistent patterns were observed for loci or populations for LD in population level tests. Similarly significant multilocus linkage disequilibrium was also detected in global tests for linkage disequilibrium (Table 3).

Allelic richness, private alleles, allele frequency and Inbreeding coefficient

Allelic richness varied among the loci from a low value of 2.34 at the locus under selection to a high value of 12.56 at neutral loci HV32 and an average value of 11.74 ± 0.57 (SD) across neutral loci (Table 2). Both *Gutierrezia* and *Solidago* feeding groups showed a large number of private alleles. 25 private alleles were observed across populations in *Gutierrezia*, and 28 private alleles in *Solidago* feeding populations (Table 4). All populations exhibited high level of inbreeding with coefficient values estimated using neutral loci ranging from 0.28 - 0.42 and with an average value of 0.33 ± 0.04 (SD) (Table 4).

Allele frequencies were estimated for the selected locus to estimate differentiation of alleles in the two host plant types. 5 alleles were observed at this locus, 3 of which were either at low frequencies or not present in many of the 14 populations. One allele which I hereafter call the "*sol* allele" was present at very high frequencies in *Solidago* populations (0.96 ± 0.03 SD). A different allele termed the "*gut* allele" was observed at high levels in all the *Gutierrezia* populations (0.68 ± 0.07 SD). *gut* and *sol* alleles were also observed in *Solidago* and *Gutierrezia* respectively at a frequency of 0.03± 0.01(SD) and 0.29± 0.07(SD) (Table 5). The *sol* allele was

fixed in two Solidago feeding populations (Ellsworth, and Marrion, Kansas). Allele frequencies at the neutral loci showed no such associative pattern for host plant groups.

Population structure: Assignment of individuals to predetermined clusters from *K*=1 to 6 using STRUCTURE 2.2 revealed that *K*=2 (LnPD value = -438.491) population clusters best explained the data in case of the selected locus. The likelihood values increase from *K*=1 to *K*=2 and then tapered off. The two clusters corresponded to *Gutierrezia* and *Solidago* populations with individuals showing distinct proportional membership to their corresponding clusters. Individuals collected from *Solidago* plants formed a *Solidago* cluster and the ones collected from *Gutierrezia* plants were correctly assigned to the other cluster (Fig 3 ; Fig 4.A & B). Unlike results using the selected locus, tests for assignment of individuals using neutral loci in STRUCTUURE 2.2 did not reveal any population structure. The LnPD value incrementally increased from K=1 to K=6 revealing no underlying structure (Fig 4.A). Assignment tests conducted in *GENALEX* 6.0 revealed results similar to that obtained from STRUCTURE 2.2. 93% of all individuals were correctly assigned to the respective population of *Gutierrezia* or *Solidago* whereas only 7% were wrongly assigned to the other population in case of selected locus. With neutral loci these numbers were 76% and 24% for self or the other population (data nort presented).

Partitioning of variance using the selected locus at different hierarchical levels in AMOVA revealed that genetic variation in populations within Kansas was structured by host plant association and not by geography. 65.34 % of the total variation was attributed to among host plant group differences, only 0.32% of the variation was observed among populations in groups

and 34.34% of the variation was attributed to within population differences (Table 6). Comparisons of only syntopic populations of *Gutierrezia* vs. *Solidago* revealed a similar structure with 57.27 % of the total variation being attributed to among host plant differences and 1.74% and 40.99% of the variance observed among populations within in groups and within population respectively. Comparisons of *Gutierrezia* in the zone of syntopy vs. other *Gutierrezia* populations as well as *Solidago* in the zone of syntopy vs. other *Solidago* populations showed no significant differences among the groups with only -0.26% and -0.97% variation attributed to among group differences respectively. 0.87 and 1.51% of variation was due to among populations within group differences. The majority of the variation was attributed to within population differences. Partitioning of variance with neutral loci revealed no significant variation among host plants with only 0.64% and 0.23% of variation being explained by among host plant differences. Similarly, among population variation explained 2.31% and 4.12% of the variation. Most of the variation was observed within populations for neutral loci (Table 6).

Mantel Test and Isolation by distance: No significant correlation between genetic distance and geographic distance was observed either with the selected locus (Z value =64.03; r=0.042; p<=0.36; geographic distance, log-transformed) or neutral loci (Z value =75.17; r=-0.1046; p<=0.80; geographic distance, log-transformed). Regression analysis of Nei's genetic distance values against the natural logarithm of geographic distance did not reveal any evidence of isolation by distance occurring in Kansas in case of neutral loci or the selected locus (Selected locus: Slope = 0.541, R^2 = 1.726e-03, Mantel P<= 0.36; Neutral loci: Slope = -0.189, R^2 = 0.0109, Mantel P<= 0.80 ; Fig 5).

F_{ST} (Genetic differentiation): Separate analyses of selected and neutral loci revealed a pattern of significant differentiation among host plants and low levels of differentiation among same host plants at the selected locus, and low level of differentiation globally at neutral loci. At the selected locus, estimates of genetic differentiation using F_{ST} values ranged from 0.42 to 0.82 in different host pairwise comparisons to 0.01 to 0.05 in same host comparisons. Comparisons of pairwise F_{ST} values among different host plant populations were similar both inside and outside the zone of syntopy. I found a slightly lower degree of differentiation in the zone of syntopy (F_{ST} 0.51-0.70) than outside the zone of syntopy (F_{ST} 0.57-0.82; Table 7). F_{ST} estimates from the neutral loci showed a different pattern of differentiation from the selected locus both among different host plants and among same host plants (F_{ST} range= 0.01-0.06) where low level of differentiation was observed in both comparisons (Appendix : Table A-1).

Gene flow: For analysis of migration rates, 14 populations were split into 4 groups (*Gutierrezia* and *Solidago* inside and outside the zone of syntopy). The program MIGRATE was run with the selected locus and neutral loci separately until values converged. Migration rates estimated using neutral loci showed low and symmetric migration into and out of the 4 population pairs irrespective of host plant type. Estimates with the locus under selection, however, revealed high migration rates among populations feeding on same host plant and low levels of migration among the different host plant populations (Fig 6). Higher migration rates were observed from populations feeding on *Gutierrezia* outside the zone of syntopy to populations feeding on *Gutierrezia* in the zone of syntopy than in the opposite direction (46.6 & 0.59). Higher migration rates were observed from *Solidago* populations of grasshoppers in the zone of syntopy to

Solidago populations outside the zone of syntopy (7.64 & 4.54) than in the opposite direction (Fig 6).

Bottlenecks: Wilcoxon sign rank tests as implemented in BOTTLENECK 1.2.02 using the two phase model (TPM) did not reveal any evidence of bottlenecks in 12 of the 14 populations studied. *Heq* (expected heterozygosity calculated from allele numbers) were not significantly different from *He* (expected heterozygozity calculated from allele frequencies). Bottlenecks were detected in 2 *Solidago* feeding populations of Marrion, Kansas and Ellsworth, Kansas (*P*= 0.01, 0.03). Both Marrion and Ellsworth are near the limit of the distribution of *Solidago* feeding populations.

Phylogeography: Neighbor-joining consensus trees were constructed based on all loci, neutral loci and the selected locus. A neighbor-joining consensus tree based on the selected locus revealed strong support for both the *Solidago*-feeding and *Gutierrezia*-feeding clusters with high bootstrap support (100% for all branches). The neighbor joining tree constructed based on neutral loci revealed weak bootstrap for the *Solidago* and *Gutierrezia* clusters even though the *Gutierrezia* and *Solidago* populations formed their own cluster except for one *Gutierrezia* population from Hoxie, Kansas and one *Solidago* population from Marrion, Kansas that clustered together. Trees constructed based on all loci clustered *Solidago* and *Gutierrezia* feeding populations separately more in line with the trees based on selected locus and showed intermediate bootstrap support (Fig 7). Clusters identified in trees constructed based on all and the selected locus correspond to the clusters identified using STRUCTURE 2.2.

Discussion

I found evidence for host-based genetic differentiation in *H. viridis* populations over relatively short distances in Kansas, and the results of the study support the role of divergent selection as the evolutionary force behind the observed differentiation pattern. Low levels of differentiation at neutral loci provide additional support for recognizing selection as the driving force structuring host plant based variation in *H. viridis*.

Role of selection and host plant based differentiation.

Significant correlation between ecological divergence, reproductive isolation and speciation has been reported in several cases (Funk et al. 2006 ; Nosil et al. 2008). Of the ecological factors, selection and adaptation to different host plants have resulted in host associated diversification in a number of organisms (Funk, 1998 ; Funk et al. 2006 ; Feder *et. al*., 2003; Sword et al. 2005 ; Nosil et al. 2008). Analysis of *H. viridis* microsatellite data revealed locus HV97 as an outlier in pairwise comparisons of populations. Divergence observed among the host plant groups was mainly due to differentiation at this locus. Loci that are outliers in multiple comparisons are likely to be influenced by divergent selection and not likely to be a result of false positives (Bonin et al. 2006 ; Nosil et al. 2008). We also observed an excess of rare alleles for microsatellite loci. An excess of rare alleles is often used as evidence for selection (Tajima, 1989 ; Payseur, 2002 ; Vigouroux et al. 2002) and divergence and thus support our interpretation. Multiple verbal models explain how speciation could result from reproductive isolation as a result of adaptive fixation of alternate alleles of genes (Muller, 1942; Dobzhansky, 1951). We identified two alleles at the selected locus occurring in varying proportions in the two host plant groups. The *sol* allele was almost fixed in *Solidago* populations with an average

frequency of 0.96 across populations. However *gut* and *sol* allele were both present in *Gutierrezia* populations at an average frequency of 0.68 (gut allele) and 0.29 (sol allele) across populations. This divergence in allele frequencies is consistent with the hypothesis of host-based selection and diversification in *H. viridis.* Analysis of the selected locus indicates that *H. viridis* populations are clustered based on host plants *Gutierrezia* and *Solidago* and exhibit high levels of divergence among them. Replicated population pairs on different hosts that are ecologically divergent and on same host that are ecologically similar allowed us to determine that host use was the likely cause for the differentiation observed at outlier/selected locus. Partitioning of genetic variance (AMOVA) across host based comparisons also indicate a host-based sub structuring in the study area. As expected in a host based differentiation scenario, AMOVA results were complemented by population assignment analyses using STRUCTURE where different clusters formed based on host plant affiliation of each population rather than by geography. These observations as well as high levels of genetic differentiation (F_{ST} values) at the selected locus and not at neutral loci in different host comparisons vs. same host comparisons further support the role of divergent selection based on host plant use in *H. viridis*.

Microsatellite data showed significant deviations from Hardy Weinberg equilibrium in overall comparisons and exhibited linkage disequilibrium (LD) in most of tests conducted. Microsatellite data reveal significant host-based clustering of individuals, validating the possibility of assortative mating on specific host plants. From the sampling schemes across the US, I found the distribution of *H. viridis* populations to be highly fragmented (Grace et al. 2009) with populations occurring in patches among a wider landscape of hosts and nonhosts. The presence of the host plant at most times did not translate to finding a population of *H. viridis.*

Assortative mating based on specific host plants and fragmented populations could result in inbreeding. High levels of inbreeding (positive F_{IS} values) and a deficit of heterozygotes was observed in the study populations. We interpret that deviations from Hardy-Weinberg could be an outcome of inbreeding and assortative mating or both. Testing for linkage disequilibrium assumes populations are in Hardy-Weinberg proportions. Observed linkage disequilibrium in several tests could also reflect a departure from Hardy-Weinberg equilibrium (Excoffier and Slatkin, 1998), which in turn was due to deficit of heterozygotes. Even though we cannot rule out physical linkage, we conclude that inbreeding rather than physical linkage between loci as the most probable reason behind the observed positive LD tests. Several studies (Machado et al. 2004 ; Ferreira et al. 2007 ; Grace et al. 2009) have reported inbreeding to cause LD. Linkage disequilibrium between markers and a deficiency of heterozygotes also point to bimodality in the study area where two parental forms thrive and dominate through reproductive isolation and thus provide additional evidence for divergent selection.

Genetic drift and differentiation.

Genetic drift could structure neutral variation across the genome and contribute to speciation in herbivorous insects as a result of the highly fragmented distribution of its hosts (Mopper, 1996). Drift could thus lead to divergent population structure as observed in the montane grasshopper *Melanoplus oregonensis* (Knowles and Richards, 2005). Based on microsatellite data, almost all variability among host plants was explained by the selected locus. Neutral loci exhibited low level of differentiation in both similar host and different host comparisons and STRUCTURE and AMOVA analysis did not show any clustering based on host plants use. Low differentiation at neutral loci thus allows us to reject the hypothesis of an

entirely allopatric origin of divergence in two refugia. This interpretation is supported by the star shaped phylogeny observed in case of mtDNA. Neutral loci and mtDNA in our study do not show much variation, potentially because less time has passed since the divergence into two host forms to allow for drift or mtDNA lineage sorting to structure genetic variation (Venketesan et al. 2007; Scheffer & Hawthorne, 2007). However, even with the neutral locus population trees clustered populations based on host plants (with very low bootstrap support) except for a pair of different host populations that were clustered together. This may suggest an initiation of genetic differentiation at neutral loci and provides limited support for the possibility that host associated selection and divergence can further promote differentiation at neutral loci even in the phase of gene flow.

Lineage diversification and population expansion.

mtDNA sequence analysis of the *Co1* gene across populations of *H. viridis* in Kansas revealed a lack of genetic structuring either based on host plant utilization or by geographic region. AMOVA analysis did not partition genetic variance among either host plant groups or among geographic regions. Testing for the endosymbiont *Wolbachia* in *H. viridis* was negative and allowed us to reject the possibility of a mitochondrial selective sweep within the limits of the *Wolbachia* genes tested. Tests conducted to identify if we have amplified mitochondrial pseudogenes returned negative results and allows us to refute that possibility (Grace et al. 2009 unpublished data). Rejecting the above possibilities, mitochondrial DNA data revealed a rapid population expansion in recent history of *H. viridis*. This pattern is supported by the recent divergence time estimate according to the molecular clock among *H. viridis* haplotypes of ~110,000 years which corresponds to the Sangamon interglacial period. The star-shaped

phylogeny indicates isolation in a single refugium followed by rapid population expansion (Avise, 1994). High haplotype diversity, low nucleotide diversity and significantly negative Tajima's D values and Fu's F_s across the distribution of *H. viridis* also support a recent rapid population growth and expansion. Haplotypes differed by only 1 to 3 substitutions indicating a recent ancestry for the different haplotypes. Of the 29 haplotypes, 27 were rare or specific to one or the other host plant. Based on mtDNA data, the 27 rare haplotypes specific to either *Solidago* (16) or *Gutierrezia* (11) may suggest an initiation of differentiation since the Sangamonian era. Host associated clusters of *Gutierrezia* and *Solidago* populations in population trees based on microsatellite data and the observed monophyly in trees based on all loci in *Gutierrezia* populations provides support for the single origin hypothesis.

Gene flow and divergence

Adaptive divergence through selection could also impede neutral gene flow among different host populations as a function of reproductive isolation between the different host forms (Nosil et al. 2008). Neutral gene flow estimates indicate symmetrical but low migration both among different host plants and same host plants. At the selected locus, low levels of gene flow were observed among different host plant groups and high levels among same host plant groups. Gene flow estimates based on the selected locus point to importance of adaptation in restricting gene flow and how it could impede neutral gene flow as a function of (a) the time since divergence and (b) the strength of reproductive isolation among the different host plant groups. Gene flow would thus decrease with increase in divergence. Neutral estimates of symmetric gene flow among and within host plant group populations could also represent historic patterns of genetic exchange and may not represent current patterns where adaptation would work to reduce

neutral gene flow. Migration rate estimates based on the selected locus indicates a progressive population expansion from *Gutierrezia* populations in the west to the zone of syntopy and then to east zone *Solidago*. Higher migration was observed from western zone *Gutierrezia* to *Gutierrezia*/*Solidago* in the zone of syntopy and from *Solidago* in zone of syntopy to eastern zone *Solidago* supporting directionality to gene flow and thus the scenario of west to east population expansion (Fig 6). The fact that bottlenecks were observed in two *Solidago* populations at the periphery of the *Solidago* population distribution in Kansas (edge populations) and no bottlenecks were observed in any *Gutierrezia* populations also support a scenario of an expansion of populations into *Solidago* from the *Gutierrezia* populations. Bottlenecks and novel selection are more likely in peripheral populations (Lesica and Allendorf, 1995) and bottlenecks often result in genetic reorganization of adapted gene complexes (Templeton, 1980). The *Solidago* populations on the eastern side of Kansas are also outside the range of the major host plant *Gutierrezia* and thus expanding populations may have been subjected to different selection pressures than within the distribution of the major host plant. From prior ecological studies (Traxler and Joern, 1999) and genetic studies using AFLP markers, there is also evidence that *Solidago* feeding is a derived state for *H. viridis* (Apple et al. 2009 submitted).

Dynamics in the zone of syntopy

Reciprocal transplant experiments conducted by Traxler and Joern (1999) found local ecological factors (sympatry) as important in promoting divergence in *H. viridis* populations from Nebraska. Populations on different hosts in the zone of syntopy in Kansas were not distinct from the populations on different hosts outside the zone of syntopy. Populations did not also show any evidence for isolation by distance at the selected locus or the neutral loci. The syntopic

populations did not cluster by geographic location as would be expected in case of a recent local host shift in that location. A local host shift or ecological speciation in the zone of syntopy in Kansas is an unlikely explanation for the pattern observed. Our study results also do not support the hypothesis of panmixia or one of full genetic differentiation. It is known that differentiation can occur even in the presence of gene flow (Servedio, 2005; Nosil et al. 2008). The fact that slightly higher variation was observed in overall comparisons of *Gutierrezia* vs. *Solidago* than in comparisons of *Gutierrezia* vs. *Solidago* in the zone of syntopy suggest the possibility of low levels of gene flow occurring in the zone of syntopy than among populations outside the syntopic zone. Restricted gene flow and high levels of genetic differentiation indicate that reproductive isolation and strong selection operate among the two host plant groups. The hypothesis of a stable zone of contact of the host races in the zone of syntopy where very limited gene flow occurs is the most likely possibility.

No-choice mating experiments conducted in our lab (Chapter 2) suggest strong reproductive isolation exists between *H. viridis* from the two different hosts. Insects collected from a host plant showed a higher preference to mate with insects collected from the same host with very few cross mating's. When insects that mated in cross matings were given a choice in the same experiment, the grasshoppers always mated with individuals from the same host. Cross mating between individuals from different hosts was higher in the zone of syntopy when compared to paired combinations of individuals from outside the zone, but these differences were not significantly different. Slightly lower levels of genetic differentiation based on microsatellite analyses within the zone of syntopy and more pairing mistakes when mating with individuals within the syntopic zone suggest that reinforcement may not be occurring in the zone

of syntopy. Considering the relatively short time since divergence, the high degree of mate preference and reproductive isolation observed among these two host plant groups across the study area, results point to the role of strong selection and adaptation to host plants driving the process of speciation in *H. viridis.* High levels of genetic differentiation, restricted gene flow, and strong pre-mating isolation suggest that this is a zone of contact but not a hybrid zone. However, studies addressing whether hybrid offspring are produced from crosses between individuals from different hosts as well as their viability have not been performed and are required to validate this.

Species growing under different selective environment neighborhoods (Brandon, 1990) could be exposed to different selection pressures in the diverse environments leading to variation among these neighborhoods. Such variation could be enhanced and selected upon in small populations. Small populations carrying a rare fitness improving allele could be subject to faster selection (Erikssson, 1998). Founding populations after rapid population expansion from a refugium are generally smaller and subject to genetic bottlenecks which could enhance divergence from the populations occupying the more suited environment (Erikssson, 1998) or host plant. Rapid selection could occur for a fitness improving allele in such cases if adaptation to a new host plant was required. An initial stage of divergence that separated *Gutierrezia* and *Solidago* feeding forms possibly in the same refugia followed by rapid population expansion and divergence is the likely explanation for the differentiation patterns observed in Kansas. However, we cannot exclude the possibility of a much recent host shift along the broad distribution of *H. viridis* in North America followed by population expansion as our study focused only on a small region of overlap of two hosts in the broad distribution of *H. viridis.* Considering the fragmented

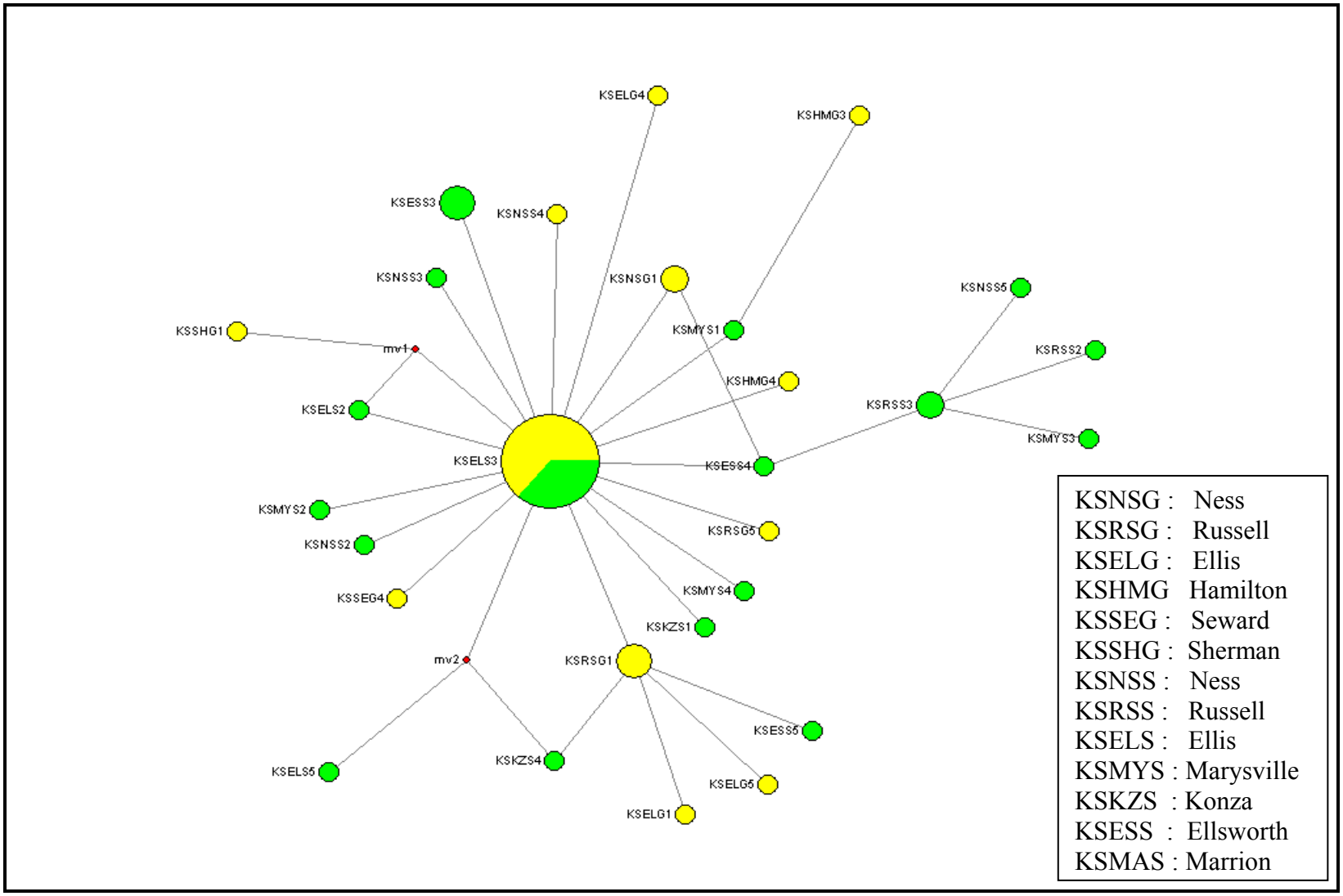
nature of *H. viridis* populations and their small effective population sizes, the founding populations possibly contained small number of individuals. Detection of bottlenecks in peripheral populations in *Solidago* and the fact that most *Solidago* populations were almost fixed for the *sol* allele may suggest costs involved in a host shift and adaptation on *Solidago* plants. Different selection pressures in different ecological environments in Kansas may have structured variations slightly differently as observed in a pair of populations in the zone of syntopy. *Sol* and *Gut* alleles observed in populations that seem to drive the divergence may have fitness values or alternatively this locus may be tightly linked to a gene or genes that affect fitness. Tradeoffs involved in host switching could reinforce selection to increase diversification among these groups. Positive correlation between divergence in host plant use and degree of mating preference provides evidence for host-related selection as the driving force in the speciation process of *H. viridis populations.* Further studies involving a larger and diverse geographical area involving additional hosts would throw light on diversification mechanism in a broader perspective.

Table 1. Mitochondrial *CO1* diversity and variability estimates in *Hesperotettix viridis populations a*cross host plants in Kansas.

Diversity Indices	**Host Plant**			***AMOVA***		**1. All Gutierrezia Vs All Solidago**				***2. Syntopic zone Gutierrezia Vs Solidago***		
					d.f.	*Sum of squares*	*Variance components*	*Percentage of variation*		*d.f.*	*Sum of squares*	*Variance components*	*Percentage of variation*
	Gutierrezia sp.	***Solidago sp.***		**Among groups**	1	1.653	0.01381	**1.48**		1	2.027	0.03294	**3.33**
n	27	28		**Among populations within groups**	11	13.559	0.09589	10.26		4	6.185	0.16298	16.47
k	12	17		**Within populations**	43	35.467	0.82481	88.26		22	17.467	0.79394	80.21
h	1.0	1.0				**3. Syntopic zone *Gutierrezia* vs Other *Gutierrezia***				**4. Syntopic zone Solidago vs other Solidago**		
					d.f.	*Sum of squares*	*Variance components*	*Percentage of variation*		*d.f.*	*Sum of squares*	*Variance components*	*Percentage of variation*
π	0.004	0.005		**Among groups**	1	1.629	0.03435	**4.42**		1	1.246	-0.00048	**-0.05**
Tajima's D	-2.1242 P< .004	-1.9611 P< .007		**Among populations within groups**	4	4.575	0.11648	14.97		5	6.110	0.05244	4.92
Fu's *F*	-27.76 P= .000	-26.92 P= .000		**Within populations**	21	13.167	0.62698	80.61		22	22.300	1.01364	95.12

1. F_{CT} = 0.015, p = 0.19 ; **2**. F_{CT} = 0.033, p = 0.31 ; **3**. F_{CT} = 0.044, p = 0.08 ; **4**. F_{CT} = -0.0045, p = 0.55 (F_{CT} : Fixation indices of among group variation)
*n, n*umber of individuals*; k,* number of haplotypes*; h,* haplotype/gene diversity*; π,* nucleotide diversity. Total *n=55* and Total *k= 29.*
Kansas *Gutierrezia*: (Ellis, Russell), Hamilton, Seward, Sherman. **Kansas *Solidago***: (Ellis, Russell) Marysville, Konza, Ellsworth, Ness.

Fig 1. Most Parsimonious median joining (MJ) network for *Hesperotettix viridis* CO1 haplotypes.



Size of the circle proportional to frequency of represented haplotypes. ; **Yellow circles** : Gutierrezia *sp*. haplotypes; **Green circles** : *Solidago sp.*

Table 2. Characteristics of seven polymorphic microsatellite loci in *Hesperotettix viridis*

Loci	***Repeat Motif***	***No. of Alleles***	H_o	H_e	A_R
Hvir54	GT-23	41	0.51*	0.95	11.71
Hvir58	GT-11	39	0.79*	0.95	11.71
Hvir73	CT-10	38	0.64*	0.94	11.09
Hvir22	CA-23	37	0.77*	0.96	12.15
Hvir97	GTT-6	5	0.24*	0.48	2.34
Hvir32	GA-23	52	0.71*	0.96	12.56
Hvir50	GT-25	29	0.36*	0.95	11.10

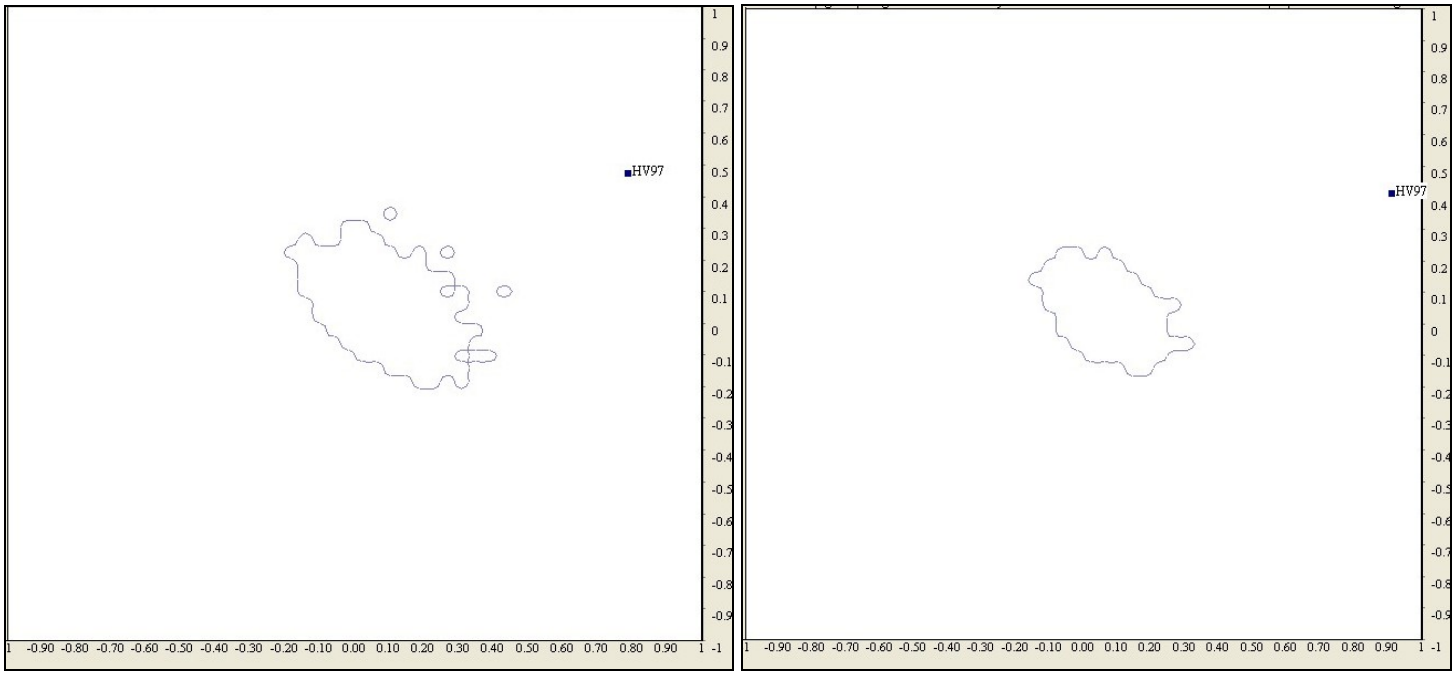
Ho, observed heterozygosity and *He*, expected heterozygosity ; A_R :Allelic richness
*Observed and expected heterozygosities are significantly different from each other (P<0.05.)

Table 3. Linkage disequilibrium between pairs of microsatellite loci in 14 populations of *Hesperotettix viridis.*

Marker	**Hvir54**	**Hvir58**	**Hvir73**	**Hvir22**	**Hvir97**	**Hvir32**	*Hvir50*
Hvir54	*	0.000	0.000	0.000	0.000	0.000	0.000
Hvir58		*	0.013	0.062	0.000	0.000	0.011
Hvir73			*	0.000	0.000	0.000	0.000
Hvir22				*	0.000	0.099	0.0043
Hvir97					*	0.000	0.000
Hvir32						*	0.000
Hvir50							*

Shaded cells indicate exact *P*-values for significant linkage disequilibrium (Adjusted *P*-value = 0.00238).

Fig 2: Output of the program DETSEL 1.0 (Vitalis *et al*. 2003), showing locus HV97 as a potential outlier (under selection) for two different host comparisons.



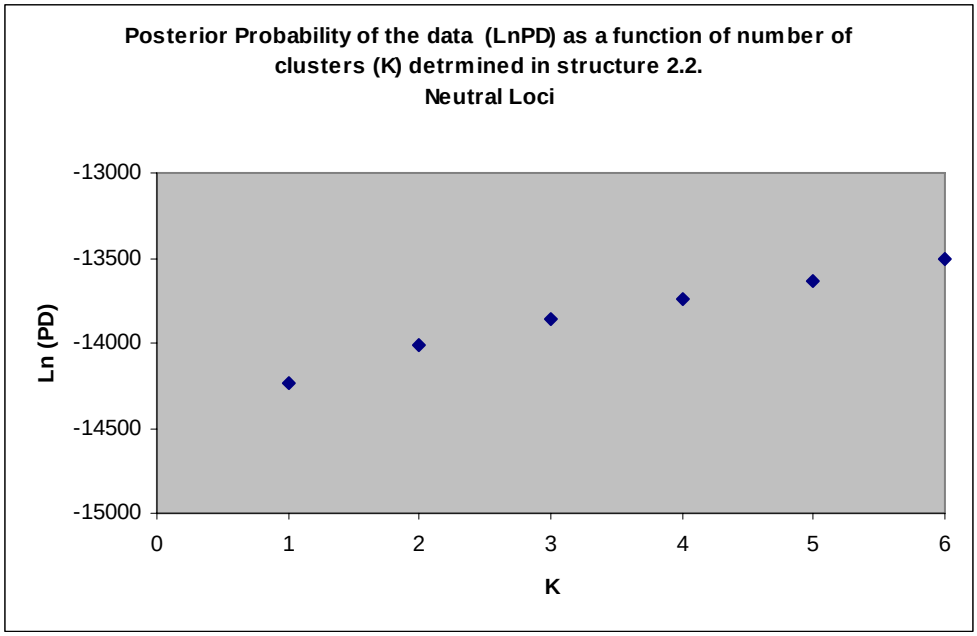
The major envelope defines the 95% high probability region derived through simulation of microsatellite data.

Russell Gut Vs Russell Sol **Hamilton Gut Vs Konza Sol**

Fig 3. Bayesian assignment of individuals from 14 Kansas populations for cluster *K=2* as in STRUCTURE 2.2.

Each bar represents an individual; Green and red colors distinguish proportional membership to different clusters. Host plant associations of individuals are indicated above each population.

Fig 4. A & B Posterior probability of the data (LnPD) as a function of the number of clusters (*K*) as determined in STRUCTURE.



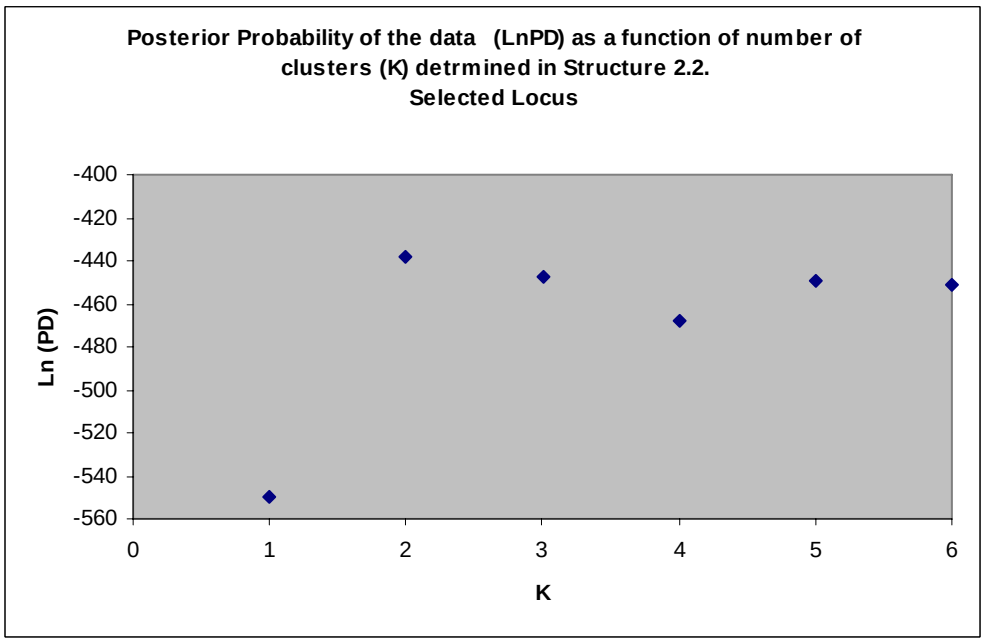
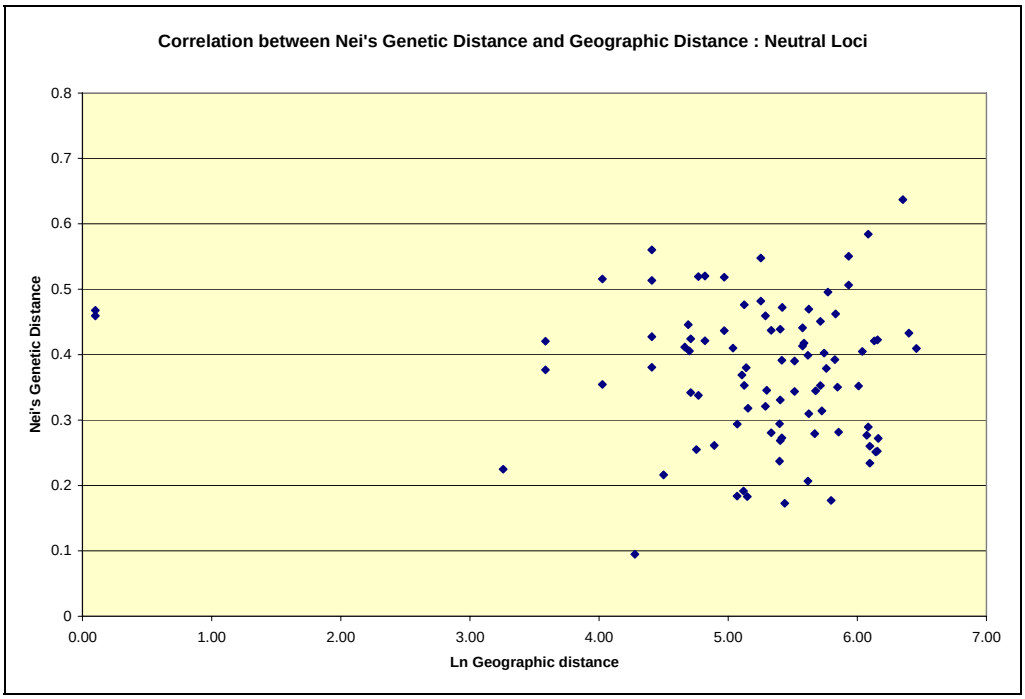
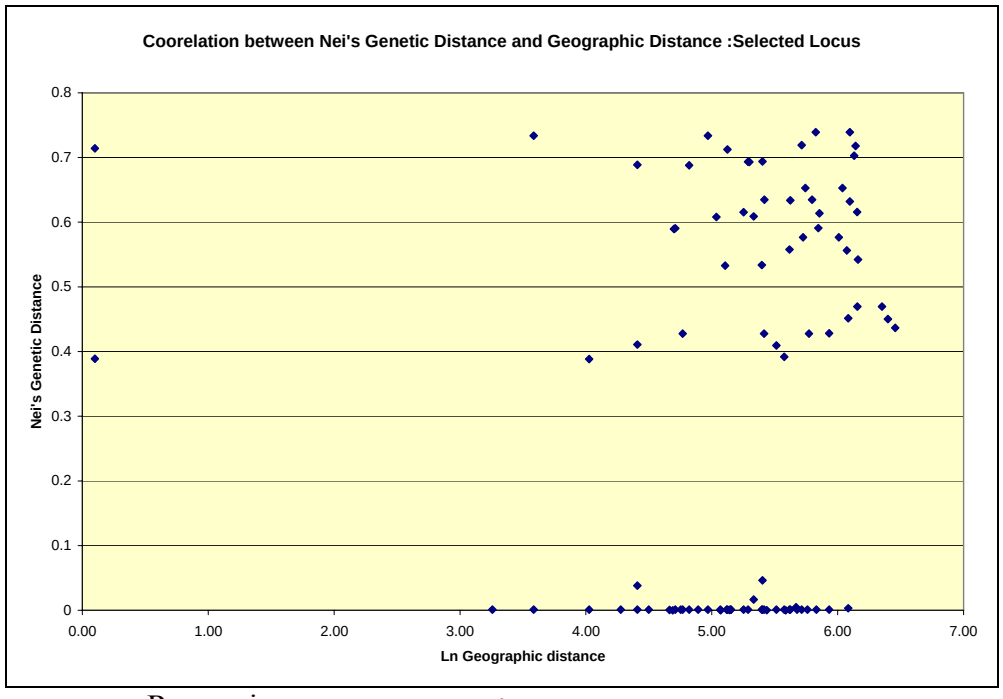


Fig 5 : Correlation between genetic and geographic distance for selected and neutral loci in 14 populations of *Hesperotettix viridis* in Kansas.

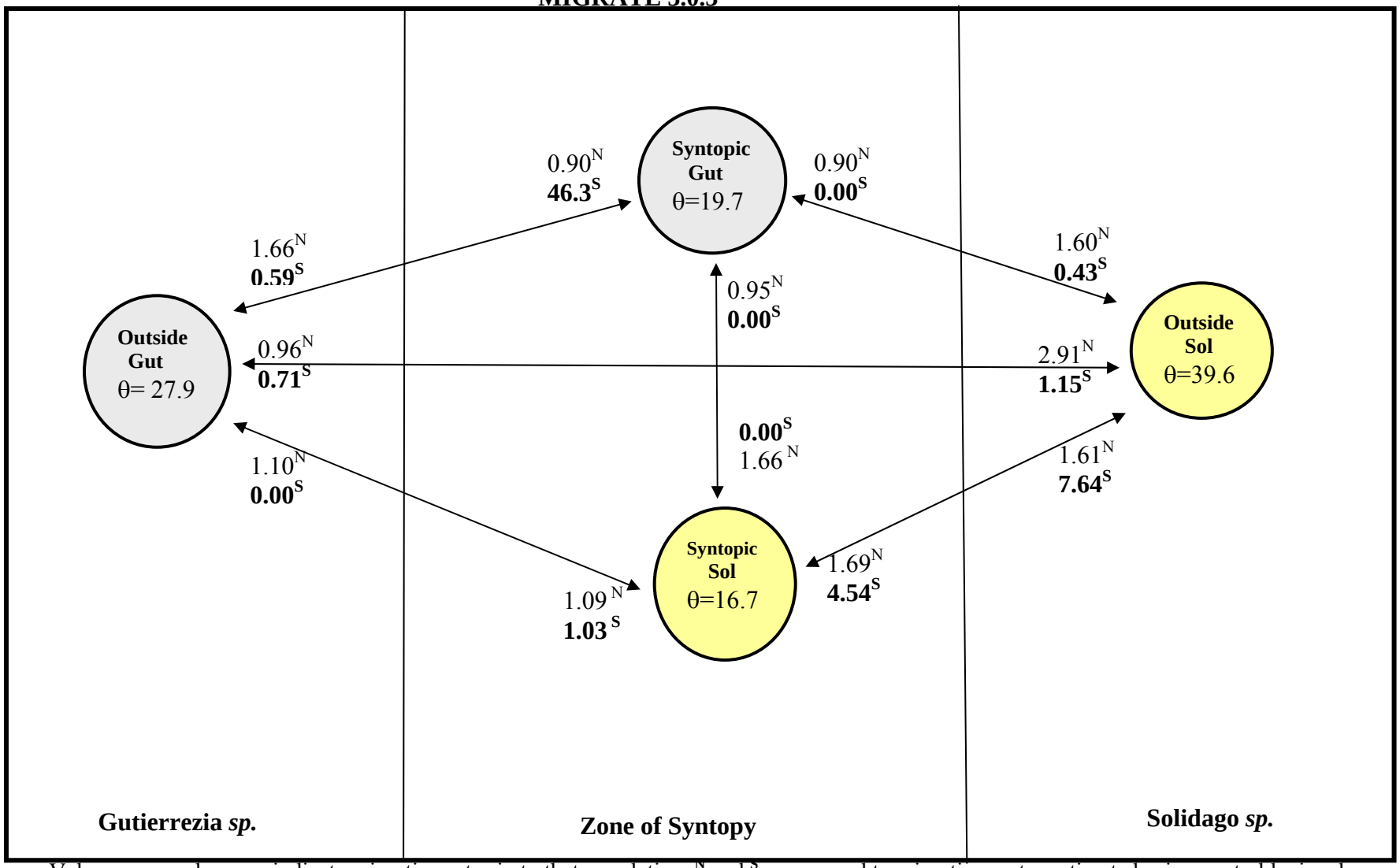


Regression Slope = -0.1897; R^2 = 0.0109 ; Mantel P<= 0.80



Regression Slope = 0.5414; R^2 = 1. 726e-03; Mantel P<= 0.36

Fig 6: Effective population sizes (θ) and migration rates in 4 populations of *Hesperotettix viridis* in Kansas estimated using MIGRATE 3.0.3



Values near each arrow indicate migration rates in to that population. ^{N} and ^{S} correspond to migrations rates estimated using neutral loci and selected loci ; θ = Effective population size estimated using neutral loci.

Table 4 : Collection location, Allelic richness, Private alleles and Inbreeding coefficient values in14 populations of *Hesperotettix viridis in Kansas.*

A_R : Allelic richness ; P_A :Private alleles; F_{IS}: Inbreeding Coefficient

Parameters	***Ellis*** *Gut.*	***Cimm.*** *Gut.*	***Sew.*** *Gut.*	***Hamil.*** *Gut.*	***Sher.*** *Gut.*	***Hoxie*** *Gut.*	***Rus.*** *Gut.*	***Ellis*** *Sol.*	***Ellsw.*** *Sol.*	***Ness*** *Sol.*	***Rus.*** *Sol.*	***Konz.*** *Sol.*	***Marr.*** *Sol.*	***Mary.*** *Sol.*	*Average*
n	26	8	19	40	34	10	27	23	36	23	25	47	12	30	
Neutral loci															
A_R	9.06	9.83	11.13	10.88	10.69	9.51	9.25	10.74	10.23	10.45	10.69	11.47	10.04	10.43	11.72
F_{IS}	0.38	0.30	0.28	0.28	0.33	0.34	0.42	0.32	0.33	0.29	0.37	0.31	0.29	0.33	0.33
Selected locus															
A_R	2.67	2	2	2.35	1.98	2.96	2.67	1.92	1	2.15	1.32	1.43	1	2.1	2.34
F_{IS}	-0.49	0.00	0.29	0.24	-0.26	0.06	-0.07	-0.03	NA	-0.05	0.00	0.66	NA	0.37	0.06
P_A by pop	2	2	2	5	4	0	0	7	1	4	2	5	0	6	Total 40
P_A by host							**Gut** 25							**Sol** 28	Total 53

Table 5. Allele Frequency of the selected locus HV97 in 14 populations of *Hesperotettix viridis*

Allele	***Ellis*** *Gut.*	***Cimm.*** *Gut.*	***Sew.*** *Gut.*	***Hamil.*** *Gut.*	***Sher.*** *Gut.*	***Hoxie*** *Gut.*	***Rus.*** *Gut.*	***Ellis*** *Sol.*	***Ellsw.*** *Sol.*	***Ness*** *Sol.*	***Rus.*** *Sol.*	***Konz.*** *Sol.*	***Marr.*** *Sol.*	**Mary.** Sol.	***All***
Sol	0.385	0.375	0.316	0.263	0.212	0.250	0.204	**0.935**	**1**	**0.913**	**0.98**	**0.968**	**1**	**0.914**	0.622
Gut	**0.558**	**0.625**	**0.684**	**0.713**	**0.788**	**0.65**	**0.741**	0.043	-	0.043	0.02	0.021	-	0.034	0.351
Other1	0.058	-	-	0.025	-	0.1	0.056	-	-	-	-	-	-	0.052	0.021
Other2	-	-	-	-	-	-	-	0.022	-	-	-	-	-	-	0.002
Other3	-	-	-	-	-	-	-	-	-	0.043	-	0.011	-	-	0.004
Total	3	2	2	3	2	3	3	3	1	3	2	3	1	3	5

Table 6. Microsatellite variance components estimated among groups and populations as per AMOVA in *Hesperotettix viridis* populations in Kansas

AMOVA		**1. All *Gutierrezia* vs. All *Solidago***							**2. Syntopic zone *Gutierrezia* vs. *Solidago***						
	d.f.	*Sum of squares*		*Variance components*		*Percentage of variation*			*d.f.*	*Sum of squares*		*Variance components*		*Percentage of variation*	
Among groups	1	13.05	82.07	0.02	0.23	0.64	**65.34**		1	9.17	20.92	0.006	0.20	0.23	**57.27**
Among populations within groups	12	71.75	2.12	0.06	0.001	2.31	**0.32**		2	17.06	0.91	0.115	0.01	4.12	**1.74**
Within populations	706	1929.0	85.08	2.73	0.12	97.05	**34.34**		198	532.96	28.79	2.69	0.14	95.66	**40.99**
		3. Syntopic zone *Gutierrezia* vs. Other *Gutierrezia*							**4. Syntopic zone *Solidago* vs. other *Solidago***						
	d.f.	*Sum of squares*		*Variance components*		*Percentage of variation*			*d.f.*	*Sum of squares*		*Variance components*		*Percentage of variation*	
Among groups	1	6.64	0.24	0.002	-0.005	0.08	**-0.26**		1	6.94	0.016	0.006	-0.0003	0.24	**-0.97**
Among populations within groups	5	27.76	1.51	0.001	0.002	2.37	**0.87**		5	30.42	0.36	0.06	0.0006	2.14	**1.51**
Within populations	321	869.59	69.91	2.71	0.22	97.55	**99.39**		385	1060.2	15.16	2.75	0.039	97.62	**99.45**

Shaded cells (A) represent selected locus and clear cells (B) represent data from the neutral loci; F_{CT}: Fixation indices of among group variation

Neutral loci : A-1. F_{CT} = 0.006, p = 0.001; A-**2**. F_{CT} = 0.002, p = 0.33 ; A-**3**. F_{CT} = 0.0007, p = 0.37; A-**4**. F_{CT} = 0.0023, p = 0.17.

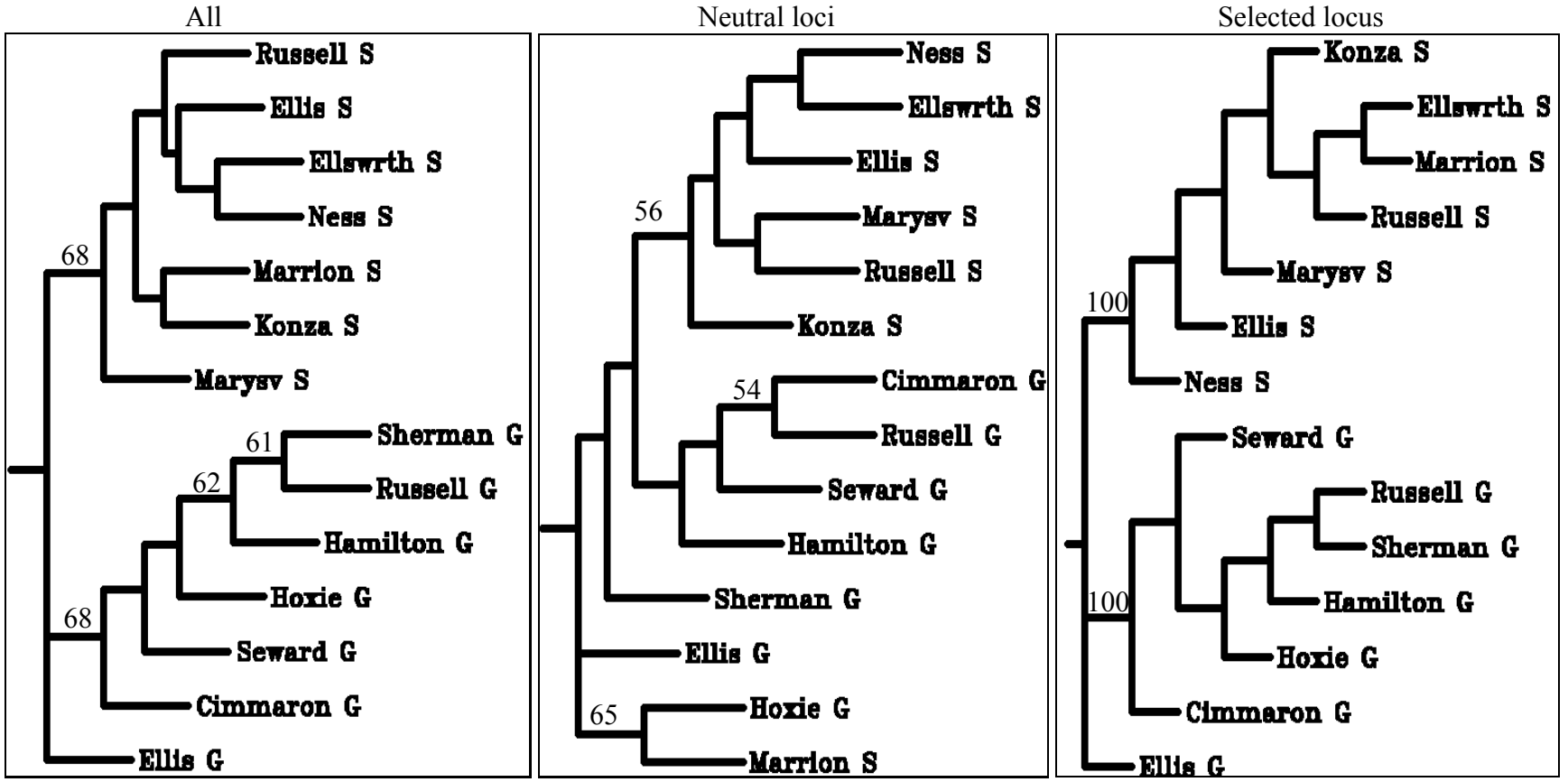
Selected Locus : B-1. F_{CT} = 0.653, p = 0.000; B-**2**. F_{CT} = 0.570, p = 0.29 ; B-**3**. F_{CT} = -0.002, p = 0.46; B-**4**. F_{CT} = -0.009, p = 0.91.

Table 7 : Pair wise genetic differentiation (F_{ST}) values within and outside the Syntopic zone at the selected and neutral loci of *Hesperotettix viridis* in Kansas

	1. Syntopic zone *Gutierrezia* vs. *Solidago* *Same Location				**2. Gutierrezia Vs other *Gutierrezia* and *Solidago* vs. other *Solidago* in Syntopic zone**			
Population	Ellis S	Russell S	Ellis S	Russell S	Ellis G	Russell G	Ellis G	Russell G
	Neutral Loci		Selected Locus		Neutral Loci		Selected Locus	
Ellis G	0.04	0.05	0.60	0.51	*	**0.05**	*	**0.05**
Russell G	0.04	0.04	0.64	0.70	**Ellis S**	**Russell S**	**Ellis S**	**Russell S**
Ellis S	*	*	*	*	*	0.03	*	0.00
Russell S	*	*	*	*	*	*	*	*

	3. Outside the Syntopic zone *Gutierrezia* vs. *Solidago* : Neutral loci					**4. Outside the Syntopic zone *Gutierrezia* vs. *Solidago* : Selected locus**				
Population	Cimarron G	Seward G	Hamilton	Sherman G	Hoxie G	Cimarron G	Seward G	Hamilton	Sherman G	Hoxie G
Elsworth S	0.04	0.02	0.03	0.03	0.03	0.82	0.75	0.70	0.79	0.81
Konza S	0.03	0.02	0.02	0.01	0.01	0.74	0.71	0.68	0.77	0.76
Marrion S	0.05	0.02	0.02	0.01	0.01	0.66	0.63	0.61	0.71	0.67
Marysville S	0.04	0.02	0.02	0.03	0.02	0.57	0.58	0.58	0.67	0.61

Fig 7. Unrooted Neighbor-joining consensus trees (50% majority rule consensus) using 1000 bootstrap replicates of Nei's genetic distance.



Letters after the location name denote host plant association; G = *Gutierrezia*, S = *Solidago*.

Chapter III

Divergent Host Plant Adaptation and Phenotypic Differentiation Drives the Evolution of Reproductive Isolation in the Grasshopper *Hesperotettix viridis* (Orthoptera: Acrididae)

Running head: Phenotypic divergence in *Hesperotettix viridis*

Tony Grace, Samantha M. Wisely, Susan J. Brown, Anthony Joern *
Division of Biology, Kansas State University, Ackert Hall, Manhattan, KS 66506

&

Elizabeth B. Maghirang and Floyd E. Dowell
Engineering and Wind Erosion Research Unit, Grain Marketing and Production Research Center, USDA, Manhattan, Kansas, USA, 66502

Keywords: *Hesperotettix viridis*, host associated divergence, allopatry, syntopy, speciation, hybrid zones, sexual isolation, reproductive isolation, phenotypic variation, feeding preference, mate choice.

*Corresponding author: Anthony Joern, Division of Biology, Kansas State University, Ackert Hall, Manhattan, KS, 66506. Phone (785) 532-6615 ; Email: ajoern@ksu.edu

This chapter is being prepared for publication submission.

Abstract

The beginning stages of lineage divergence can be difficult to detect, as correlations between altered genotypic and phenotypic attributes are often weak early in the process. Shifts in host plant use and divergence in mating signals can lead to sexual isolation and ultimately speciation. To understand the process, it is important to recognize the initial barriers that reduced contact between populations, the evolutionary factors leading to adaptation if it occurs, and the contributions of such factors and traits to better understand lineage divergence. The grasshopper *Hesperotettix viridis* is an oligophagous species feeding on host plants primarily in the genus *Gutierrezia* and *Solidago* in Kansas. Genetic studies conducted in our lab suggest that selection contributes to genetic differentiation in response to host use. Here I present results examining the host preferences, degree of differentiation in mate choice, and divergence in cuticular morphology using near infrared spectroscopy. I compared the divergence in these traits on different host plants in *H. viridis populations* in Kansas in zones of contact and in populations with access to only one host species to identify the nature and origins of putative initial barriers that isolate populations in the formative stages of divergence. Significant host-based preference of individuals from the two host plant groups was detected in host paired-feeding preference studies. No-choice mate selection experiments reveal preference for individuals collected from the same host species independent of location, and little mating observed between individuals from different host species; female mate choice tests between males from the two host species resulted in 100% fidelity with respect to host use. Significant differentiation in color and cuticular composition of individuals from different host plants was also observed. A correlation between host choice and mate

choice and phenotypic divergence resulted. No evidence for reinforcement or character displacement in the zone of contact was detected. Results of our study indicate that divergent selection for host plant use promotes sexual isolation in this species.

Introduction

Understanding mechanisms of speciation for insect herbivores have centered around host plant preferences and host shifts (Feder et al. 2003; Egan et al. *2008*; Nosil et al. 2008). Because the evolution of behavioral differences in feeding can promote speciation in animals by reducing mating between individuals in diverging populations, local divergence can evolve through combined influences of feeding and mating preferences associated with host plant use (Feder and Forbes, 2008 ; Funk and Nosil, 2008). Alternately, individuals feeding on different hosts may exhibit phenotypic plasticity, the ability of a single genotype to exhibit different phenotypes in different environments that vary in physiology morphology and behavior in response to environmental cues (Agarwal, 2001; Gorur et al. 2007). While phenotypic plasticity can be adaptive with many ecological benefits, it does not necessarily promote speciation. Plasticity in host choice may lead to mate choice and divergence in herbivores using different host plants (Gorur, 2005), but the presence of phenotypic plasticity in life history traits alone does not contribute to speciation and genotypic variations in phenotypic plasticity among host plant groups is essential factor in deciding if divergence occurs (Gorur et al. 2007). Recently, studies have focused explicitly on the role of natural selection and ecological factors in reproductive isolation and speciation (Funk, 1998; Egan et al. 2008; Nosil et al. 2008). Strong associations develop between an insect herbivore and its host plant, both as food and primary habitat, accentuating the likelihood that local adaptation will occur with respect to differences in host plant use, and thus beginning a chain of events leading to phenotypic and genetic differentiation (Mopper and Strauss, 1998; Berlocher and Feder, 2002 ; Gavrilets and Vose, 2009).

Diet selection in insect herbivores often varies among locations in response to ecological, physiological, behavioral and evolutionary forces (Berenbaum and Zangerl, 1998). Constraints from physiological/biochemical trade-offs that limit performance on other hosts are one possibility. Such trade-offs would reinforce the evolution of host specialization (Joshi and Thompson, 1995; Berenbaum and Zangerl, 1998). Even in sympatry, insect herbivore populations often exhibit fine scale genetic differentiation in association with particular host plant species (Mopper, 1996). Positive correlations between host use and performance on existing host plants could reinforce local adaptation to the particular hosts and lead to host races (Joshi and Thompson, 1995., Thomas and Singer, 1998). This intense association with the host plant makes them susceptible to disruptive selection if and when they change hosts (Nosil, 2005; Nosil, 2007; Berlocher and Feder 2002; Mopper and Strauss 1998).

Speciation in herbivorous insects could result from geographic isolation (allopatric speciation) or from interactions among local factors (such as host availability) leading to host shifts followed by local adaptation and strong selection. Phytophagous insects that feed on specific host plants provide an ideal system to understand the geographic and ecological and other factors that promote divergence due to their intimate relationship to their host plants. Feeding patterns observed may either represent choice or phenotypic differences with no genetic basis or represent ancient or recent host shifts leading to adaptation and divergence on specific host plants. The latter case is expected for host races exhibiting host fidelity with correlations between host choice and mate choice (Dre's and Mallet, 2002).

Divergence in host plant preference begins the process of pre-mating isolation (assortative mating from mate choice) through habitat isolation and sexual isolation (Via et al. 1999 ; Nosil et al. 2006), a process that can be enhanced by divergence in phenotypic traits (like color, cuticular composition, and behavioral traits). In cases where divergent selection drives speciation, populations on different host plants at the same location may exhibit a greater degree of reproductive isolation than populations on the same hosts in allopatry (Funk, 1998). If geographic isolation is the only factor that contributes to divergence, one expects a strong isolation by distance signature. In such cases, drift as well as selection could structure variation over time because of isolation and the reduced opportunity for gene flow among populations. Populations that come into contact after evolving in allopatric isolation could develop incompatibilities that become reinforced/diluted when they come back in contact, the degree depending on the strength of discrimination mechanisms that evolved when populations were separated. Some mechanisms acting to reinforce mating isolation include selection against hybrids, lower fertility and higher mortality in inter-specific mating (Servedio, 2005). After contact, reinforcement could further promote differentiation as isolation mechanisms are strengthened. Because selection for reinforcement occurs only in sympatry and not in allopatry, greater selection for mating preference should be detected in areas of secondary contact in comparison to allopatric areas. This has been demonstrated in several studies involving mate choice and hybrid viability assays including walking sticks *Timema cristinae* and *Drosophila pseudoobscura* (Noor, 1995 ; Nosil et al. 2002). A prediction of reinforcement is that geographically coincident populations should exhibit greater strength in mate preference than that observed in geographically separated populations.

Feeding on different hosts may affect other organismal traits that can feedback and affect divergence processes. For example, diet composition can influence the abundance and composition of hydrocarbons in the cuticle (Chapman et al. 2000), a factor in phytophagous insects that feed on multiple hosts. Cuticular chemicals are believed to play a role in communication during courtship, and increase behavioral isolation of host races. They may thus contribute to the process of speciation if divergence of such signals leads to non-random mating between populations (Neems and Butlin, 1994). However phenotypic differences alone do not contribute to speciation and genetic factors underlying phenotypic differences play a major role as to if divergence occurs (Gorur et al. 2007). Indeed, cuticular hydrocarbons have been used in taxonomic studies to differentiate among species in Isoptera (Haverty et al. 1988), Diptera (Phillips et al. 1990), Coleoptera (Lockey and Metcalfe, 1988) and Orthoptera (Chapman et al. 1995, 2000). Variation in cuticular hydrocarbons has also been reported across a hybrid zone among sub-species of the meadow grasshopper *Chorthippus parallelus* (Neems and Butlin, 1994). Phenotypic divergence of some morphological traits can thus aid in sexual isolation and reinforce divergence or it can occur as an after effect with increasing pre-mating isolation.

Hesperotettix viridis (Thomas) is an oligophagous grasshopper species that feeds on composites (Asteraceae) (Joern, 1979, 1985, Traxler and Joern, 1999). Individuals spent most of their life and activities such as feeding, thermoregulation and mating on the same host plant (Parker 1983, Pfadt 1994, Traxler and Joern, 1999). Adults have functional wings but are poor fliers. Eggs are laid as pods in soil beneath the host plant

(Traxler and Joern, 1999). In Kansas, I observed two distinct *H.* viridis phenotypes that feed on the host plants in the composite genera *Gutierrezia* and *Solidago.* The distribution of *H. viridis* in Kansas is such that eastern populations encounter only *Solidago* and western populations encounter only *Gutierrezia*. Some *H. viridis* populations occur at geographically intermediate locations where both host plant populations co-occur syntopically.

Chapter 1 of the thesis document genetically-based host race diversification in *H. viridis* populations in Kansas. A host race is defined as populations that use different host plants showing differentiation and consist of individuals that exhibit strong host fidelity, coexist in sympatry in at least part of their range, are genetically differentiated at more than one locus, and are spatially and temporally replicable (Dre`s and Mallet 2002). This suggests that individuals should discriminate according to host plant use and preferentially mate with their own host race. Reciprocal transplant experiments under field conditions in Nebraska revealed divergence in ecological performance and host preference by *H. viridis* in sympatric populations, including decreased fitness in mismatched host plant pairings (Traxler and Joern 1999). Such responses to host plant use at the physiological level could act to isolate host-based populations and reinforce the evolution of diet specialization.

In this study, I tested the relative contributions of host plant preference and sexual isolation in ecological divergence and genetic divergence of *H. viridis* host populations in Kansas. Significant preferences for the ”home” host would be expected, if host plant use

drives ecological divergence. Character displacement is expected to occur among the host plant groups if higher divergence in feeding preference among the groups is observed in regions where they co-occur than locations where the distributions do not overlap. If adaptation to specific host plants drives sexual isolation, increased sexual isolation should occur more strongly among ecologically different pairs of populations than among ecologically similar populations. The presence of host-based assortative mating in *H. viridis* would suggest that ecological divergence is reinforced by pre-mating mechanisms. Higher sexual isolation is expected to occur in regions of syntopy (regions in which both species co-occur within cruising range) than in allopatric regions if the observed pattern of sexual isolation is a result of reinforcement. Differentiation in phenotypic characters can also affect genetic differentiation (Weiss et al. 2006). To investigate the existence of possible phenotypic differentiation associated with genetic differentiation, I also characterized variation in cuticular characteristics and color morphology among host populations of *H. viridis* using near-infra red (NIR) spectroscopy. Variations in cuticular composition are often associated with levels of assortative mating in the grasshopper *Chorthippus parallelus* and divergence in such mating signals could lead to pre-mating reproductive isolation (Tregenza et al. 2000). Variation in phenotypic characteristics that allow mate discrimination could speed up the tempo of genetic differentiation or result as an outcome of differentiation. It is important to ascertain phenotypic differentiation, if any exists among the host variants, to better understand their role in initial events that could ultimately lead to speciation. I hypothesize that if variation in visual or cuticular characters are present in *H. viridis*, they would result in unique visible or NIR spectra for the different host variants. Ultimately, to better understand host race formation and

speciation, one must recognize opportunities for reproductive barriers that reduced contact initially between populations, the evolutionary factors behind those Coyne and Orr, 2004) and the contributions of various factors as host based selection, phenotypic variations etc. that promote divergence.

Materials and Methods

Sample collection

The grasshopper *H. viridis* is patchily distributed in Kansas. Populations of *H. viridis* were collected from areas where potential host plants of *Gutierrezia sp.* and *Solidago sp.* were known to coexist (zone of syntopy) and from areas with pure stands of only one of the two host plants (*Gutierrezia sp.* on western side and *Solidago sp* eastern side of the zone of syntopy. 40-60 individuals from each of the 8 populations (4 from *Gutierrezia* and 4 from *Solidago* populations encompassing East, West, Central regions) were collected from each site in Kansas in 2008. Collection details, location and host plant association of the populations are given in Table 4. *H. viridis* individuals were collected from their respective host plants by hand. The activity of the respective insect on its host plant (feeding, mating, or resting on host plant) was recorded to ascertain host plant association. Each individual was collected alive, brought back to the lab and reared until adult emergence. Insects were collected fairly early in the season and most individuals were later stage nymphs or recently emerged adults. Sexes were kept isolated in separate cages and were fed with foliage of plants on which they were collected.

Feeding trials. Choice assays were conducted where insects from 8 populations (n=299) were provided simultaneously a choice of both *Gutierrezia* and *Solidago* plants. Individual insects were placed in a 18x13x10cm ventilated clear plastic arena with one ~10cm cutting from each of the two host plants, the bottom of which was placed in small plastic vials filled with water to keep the plants fresh for the duration of the feeding trials. Feeding assays were initiated in the morning and conducted for 8 hours following which the insects were left overnight in the container. Insects were fasted for 24 hours prior to

commencing feeding trials. Data were recorded as observed feeding activity every 30 minutes for 8 hours. Each individual was only used once in the feeding trials. Most insects fed on either *Gutierrezia* or *Solidago* alone. But some individuals (<4%) fed on one host plant during one observation interval and on the alternate host in several observation intervals. The observations were summarized as 0 or 1 for each insect that fed on *Gutierrezia* and *Solidago,* respectively. Observations for the few insects that were observed feeding on multiple hosts were also summarized as 0 and 1 based on majority time observed feeding on a particular host.

Mating Trials. (a) No-choice mating trials were conducted for the 8 populations by placing a combination of one male and one female from the same host or alternate host into a clear plastic arena as in the feeding study. Observed mating activity was recorded every half an hour for 8 hours after which the insects were left in the plastic arena overnight. Each individual was only used once in the mating trials. The observations were summarized as 0 for not-mated and 1 for each insect pair that mated. The experiment was separated into 2 groups of four populations each and all pair wise combinations of mating trials were conducted. Ten pairwise combinations of mating trials were conducted for each group, of which 4 included combinations of individuals from different hosts, and 6 trials were same-host combinations within a group. Each pairwise combination of mating trials was replicated 16 times of which for the first 8 trials, males collected from *Gutierrezia* were paired with females from *Solidago* and the next 8 were females *Gutierrezia* with males of the *Solidago* (for different host trials). A total of 320 mating trials (16 replicates x 20 crosses) were conducted.

(b) Paired-choice mating trials were conducted the next day after no-choice experiments for (1) trials where insects mated with the other host plant form (cross matings), and (2) trials where insects did not mate when the mate was from the alternate host. For both trials, an equal number of mating pairs were used (n=18 each for 1 & 2). A male from the same host plant as the female was added (in addition to the male already present from the different host) to the clear plastic arena and observations were recorded as for the no-choice trials. In trial one, I intended to test mating propensity when a choice was available to ascertain if an insect from a specific host would prefer a mate from the same host plant. In trial 2, I tested for female receptivity as a factor in mating decisions (in no choice assays) by providing the insects that did not mate with a choice. This combination was essentially a control to determine that a female that did not mate actually would if the proper male was present.

Near Infrared Measurements. Near infrared spectra (NIR) can be used to compare cuticular absorbance at specific wavelengths between individuals from different hosts as an indication of variability in structural components of the cuticle, and to measure variation in visual characteristics in the two host forms. Here, near-infrared spectra of individual insects from each of the 8 populations were collected. Spectra (450-2500nm) were collected from the thorax of each grasshopper using a ASD QualitySpecPro Benchtop Spectrophotometer (Analytical Spectral Devices, CO, USA). The NIR spectra were recorded at 2nm intervals as absorbance (log1/Reflectance) mode in the 450nm-2500 nm range using the RS3 software. Individuals were positioned dorsal side down and the samples were illuminated with light from the bifurcated fiber optic reflectance probe (4.8mm in diameter) positioned 5mm from the bottom of the viewing area. Considering

the size difference between the sexes, the probe was directed upwards to maintain the 5mm distance of the probe to the viewing area of ASD trumpet accessory across samples. Because males in *H. viridis* are smaller than females, males and females were analyzed separately to account for any sex based variability, and the results were weighted to account for unequal sample numbers from each host. The instrument was optimized and a new baseline was determined at the beginning of each set of population samples.

Statistical Analyses

Feeding trials. Tests were conducted to determine whether individuals collected from *Gutierrezia* exhibited different host preferences than did individuals derived from *Solidago* using a chi-squared analysis. All *Gutierrezia* populations and *Solidago* populations were grouped together irrespective of the geographical location from which they were collected and analyzed separately to assess any differences. I also analyzed separately populations from allopatric and syntopic areas to detect the presence of any differences in host preference between populations in allopatry and syntopy. Tests were also done to determine whether the strength of host preferences differed between populations in allopatry vs. syntopy using logistic regression by testing for an interaction between host plant use and geography. Presence of a significant interaction term would indicate that feeding preferences differ in allopatry and syntopy. I also tested for character displacement in the zone of syntopy using the procedure described in Nosil et al. (2006) to estimate divergence for populations. Divergence for 2 pairs of allopatric and 2 pairs of syntopic population pairs was calculated as the percent of individuals from the

Gutierrezia populations preferring *Gutierrezia*- % individuals from *Solidago* preferring *Gutierrezia.* All analysis was done in *R* ver. 2.9.0 (R Development Core Team).

Mating Trials. I estimated a reproductive isolation index I_{PSI} (Rolan-Alvarez & Caballero, 2000), which captures the intensity of mating isolation among the host plant groups. The significance value and standard deviations for this index were estimated based on resampling the data 10,000 times. Copulation frequencies were analyzed using logistic regression in a model that calculated the probability of copulation based on male host, female host and an interaction term. A significant interaction between the male and female host indicates that reproductive isolation among the two host forms is present. The significance of the model was ascertained from Akaike information criterion (AIC) values (Burnham and Anderson, 1998), which is essentially equivalent to the cross validation criterion. The lowest AIC value corresponds to the best model (maximum log likelihood for the fitted model+ number of covariates including the intercept). For the different host combinations, I also tested whether reinforcement occurs in a model where copulation frequency was examined with males from 8 populations correlated to the male host, female host and allopatry (i.e., I tested whether females are from an allopatric or syntopic population). All analyses were performed with a reduced regression model *R* ver. 2.9.0 (R Development Core Team) using backward elimination until the best model with the lowest AIC value was attained. Mean mating probabilities and 95% confidence intervals of different-host and same-host crosses were estimated by logistic ANOVA in the GENMOD procedure in SAS ver. 9.1.3 (SAS Inst.).

Near Infrared Spectroscopy: Data validation and analysis: Near-infrared spectral data were validated and analyzed using partial least squares (PLS) regression analysis with GRAMS32 software (Thermo Galactic Industries, NH, USA) for two-way comparisons between host plant use. Spectral scores (independent variable) were regressed against the host plant group (dependent variable) in the PLS regression. One *Gutierrezia* population from Sherman and one *Solidago* population from Marysville were used to develop the model and the remaining 6 population samples were used to validate the resulting model. For validation, the populations from Sherman and Marysville were treated as known samples, and all samples were subjected to a cut-off "rejection threshold" of 1.5. Any individuals collected on *Gutierrezia* with a partial coefficient >1.5 were considered misclassified, while any individuals collected on *Solidago* with values < 1.5 were considered as misclassifications. Wavelengths important in differentiating the potential host races were determined by interpreting the spectral results and validating those interpretations using partial least square regression coefficient values. The accuracy of classification was interpreted from the percent correct classification, the coefficient of determination and the standard error of cross validation. Three different sets of partial least square regression analyses were performed: (1) All *Gutierrezia* vs. all *Solidago* populations: G1 + G2 + G3 + G4 versus S1 + S2 + S3 + S4; (2) *Gutierrezia* vs *Solidago* populations outside the zone of syntopy: G1+ G2 versus S1 + S2; and (3) *Gutierrezia* vs *Solidago* populations within the zone of syntopy: G3 + G4 versus S3 + S4 (Table 6). Analyses were performed for 3 different wavelengths to identify regions of spectrum most useful in classification (1) 450 to 2250nm, (2) 750 to 2250nm, and (3) 450 to 750 nm.

Results

Host Plant Preferences: Populations of *H. viridis* collected from *Gutierrezia* and *Solidago* showed significant host preferences, where both groups fed on the primary host from which they were collected (Both $P < 0.001$; Table 1). Insects collected on *Gutierrezia* showed a higher preference for *Gutierrezia* than did individuals collected from *Solidago* showed for *Solidago* plants (Table 1B). 96% and 84% of the insects collected from *Gutierrezia* and *Solidago* fed on *Gutierrezia* and *Solidago*, respectively. The frequency of individuals feeding on host plants at each location is shown in Fig 1. I analyzed allopatric and syntopic populations separately to determine if feeding preference divergence differed in the two scenarios and whether character displacement occurred in the zone of syntopy. In allopatry, individuals that fed on *Gutierrezia* showed a higher preference for *Gutierrezia* than did individuals that fed on allopatric *Solidago* and *vice versa* ($P < 0.001$; Table 1). Similarly, individuals from populations on *Gutierrezia* in syntopy showed a higher preference for *Gutierrezia* than was observed for synoptic populations collected from *Solidago*. Again, populations on *Gutierrezia* in areas of syntopy were less willing to feed on the alternate host than was observed for populations on *Solidago;* these differences were not statistically significant. I also found a significant preference for the home host from which the population was collected in all 8 populations (Table 1; all $P < 0.01$). One population from Ellis feeding on *Solidago* showed less divergence in preference than the other populations on either of the hosts.

Logistic regression analysis did not reveal any significant interactions between host use*geography (AIC = 180; P =0.47). Thus, host preference within a plant host group did not differ significantly between individuals from allopatric vs. syntopic

geographic regions. Tests for character displacement revealed that for syntopic population's pairs, the mean of differences was 74.5% ; SD=7.7 and for the 2 pairs of allopatric populations the mean difference was 93.2% ; SD=5.3. (P = 0.03).Thus divergence was lower in syntopic regions than in allopatry, a trend opposite of that expected if character displacement was occurring.

Mating Tests. No-choice mating trials in 20 paired combinations among all populations revealed strong host-based sexual isolation among the two host plant groups feeding on *Gutierrezia* and *Solidago*. The proportion of mating trials in which copulation occurred is given in Figure 2. Mean mating probabilities and 95% confidence intervals of different-host and same-host crosses are shown in Table 3. Irrespective of the geographic location, insects in trials from a host plant population always mated at higher proportions with individuals from the same host plant than in trials where a mate from the other host plant group was provided. The magnitude of sexual isolation varied from 0.46 ± 0.20SD in Ellis *Gut* vs. *Sol* (syntopic) to 0.89 ± 0.11SD in Marysville *Sol* vs. Sherman *Gut* (allopatric) comparisons with an average value of 0.75 ± 0.13(SD) across all different host comparisons (Table 2). Within the same-host comparisons, sexual isolation index values were very low and less than 0.036 in all comparisons, indicating little to no sexual isolation among individuals from the same hosts. A significant interaction between male-host and female-host was observed in the logistic regression model also supporting strong host-associated divergence in mating frequencies (AIC value =144; 1 df; P <0.0001). A model that tested for copulation frequencies of males with females (from either allopatry or syntopy) revealed a negative correlation between mating frequencies and geographic distance. (AIC value = 99.11; P-value = 0.041). Thus the hypothesis that reinforcement

exists is not supported. Additionally, based on the mating frequency data, it is clear that more mating pairs occurred in trials using insects from within the zone of syntopy than in trials involving allopatric pairs.

Results from mate choice trials for insects that mated from different hosts (in no choice trials) show that in all but one trial, mating occurred with an individual from the same host plant. In one trial the female did not mate with males from either the same host or alternate host. Similarly, for trials where insects did not mate originally when presented with an individual from the alternate host plant group, individuals mated 100% of the time when a choice from the same host was provided. Combined, these two results suggest strong assortative mating based on host plant form.

Cuticular NIR Profiles: Cross-validation results of NIR spectra reveal that in comparisons of all individuals collected from *Gutierrezia* vs. all individuals from *Solidago*, 96.4% of individuals were correctly classified into the correct host plant cluster. In comparisons among populations outside the zone of syntopy, 99.4% of all individuals were correctly differentiated into respective clusters of those that fed on *Gutierrezia* vs. *Solidago*. Within the zone of syntopy, 96.7% of individuals were correctly assigned in comparisons of those that fed on *Gutierrezia* vs. *Solidago*. The coefficient of determination (R^2) for the calibration model of all *Gutierrezia* vs. all *Solidago* was 0.80. R^2 for comparisons of *Gutierrezia* vs. *Solidago* outside the zone of syntopy was 0.84, and the value of R^2 within the zone of syntopy comparisons of *Gutierrezia* vs. *Solidago* was 0.81. Within and outside the zone of syntopy, similar patterns of classification success and strength of correlation were observed (Table 6). Partial least square regression analysis was used to assess which wavelengths were most

useful in discriminating *H. viridis* populations. Eight wavelengths were useful in identifying *H. viridis* populations in the three comparisons among models. Three wavelengths of 510nm, 690nm and 722nm were in the visible region of the spectrum and 800nm, 1050nm, 1360nm, 1820nm and 1900nm were in the NIR region (Table 6). Of all the wavelengths useful in discriminating among *H. viridis* populations according to host plant use and based on percent correct classification, the visible region of the spectrum representing color differences was most important. Other important regions useful in identification represent the CH3 combination overtone (1360nm), O-H stretch/C-O stretch 2nd overtone combination corresponding to cellulose in plants (1820nm), and C=O stretch 2nd overtone corresponding to (–Co2H) carboxylic group (1900nm) (Shenk et al. 1992). Average spectra of all *Gutierrezia* and all *Solidago* populations of *H. viridis* are shown in Figure 3.

The calibration dataset at different wavelengths (450-750nm, 750-2250nm and 450-2250nm) using the Sherman (G1) Marysville (S1) calibration models permitted us to infer the relative importance of different regions of the spectrum in distinguishing between the host races. The visual region of the spectrum was the most important region enabling identification between the host races with 97.9% and 98.1% correct classification into *Gutierrezia* and *Solidago* clusters. 750-2500nm wave length could correctly identify the individuals of *Gutierrezia* and *Solidago* populations with 81.4% and 98.0% accuracy (Table 5). When the entire region of the spectrum was considered (450-2500nm), accuracy was 95.2% and 98.1% for *Gutierrezia* and *Solidago* individuals, respectively.

Discussion

Early stages of speciation are characterized by the initiation of genetic differentiation in response to reduction in gene flow, genetic drift or natural selection, something that may occur when insect herbivores shift onto different host species. In this study, host associated feeding divergence and assortative mating based on host plant use by *H.* viridis was observed for two primary host species in the genera *Solidago* and *Gutierrezia*. Results also indicate that *H. viridis* populations on different host plants have diverged in color and cuticular characteristics, indicating that divergence in feeding preference can promote population differentiation. By comparing responses of *H. viridis* populations encountering only one of two possible hosts with populations with access to both hosts, I determined the role of host plant use in initial divergence, the effect of habitat preference on sexual isolation and phenotypic divergence in color and cuticular characteristics among host forms. Feeding preferences on different hosts and mating isolation were observed in all populations and comparisons within Kansas. Divergence in feeding preference was not significantly different in syntopic populations vs. allopatric populations. However more insects from within the zone of syntopy were willing to feed on the alternate host than were allopatric populations presented with alternate hosts. A slightly negative impact on mating with geographic distance was observed in no-choice trials, where mating frequencies decreased as distance between sites increased. More pairings occurred in the zone of syntopy than among geographically isolated populations. In choice experiments, all but one individual (trial) that copulated with a mate from the different host in no-choice trials mated back with a mate from its own host plant when a choice was provided. Also 100% copulation frequency in female receptivity trials (choice

tests) proved that receptivity was not a factor when mating between different host forms was not observed. Mate choice experiments provide the strongest evidence for assortative mating based on host plants, which further emphasizes the importance of host plant use in mating decisions. Phenotypic differentiation among host forms in color and cuticular characteristics is also supported by our study.

Feeding preference and character displacement

Feeding preference often varies among locations in response to ecological, physiological, behavioral and evolutionary forces (Berenbaum and Zangerl, 1998). Constraints from tradeoffs in host switching limiting performance on other hosts would reinforce the evolution of host specialization (Joshi and Thompson, 1995., Berenbaum and Zangerl, 1998). It is known that strong selection against host switching could lead to the evolution of preference for alternate hosts (Nosil et al. 2006). I observed strong feeding preferences for "home" hosts in all of the eight Kansas populations studied. Divergence in feeding preference was not significantly different between syntopic populations vs. allopatric populations. Instead, more insects fed on alternate hosts within the zone of syntopy than in allopatric populations, a result contrary to what is expected if character displacement is occurring in the zone of syntopy. When character displacement occurs, divergence among two groups should be accentuated in regions where they co-occur than where the distributions do not overlap (Brown and Wilson, 1956). The specific test for character displacement also revealed lower divergence in syntopic populations than in allopatric populations. This is conclusive evidence that character displacement did not take place as lower mean divergence values were observed in the

zone of syntopy. In addition, lower divergence suggests that limited gene flow may be occurring between populations in the syntopic zone for both host species than in allopatry where selection regimes for the host plant may be stronger. However, strong divergence in feeding preference observed in the zone of syntopy and in allopatry indicates that use of different hosts is the initial mechanism promoting divergence.

Reinforcement in the zone of syntopy

Feeding specialization may set the stage for reduction in mating between populations feeding on alternate hosts (Via, 1999). Reinforcement occurs when there is a higher degree of mating divergence when populations on different hosts come into contact than in allopatry, reducing the likelihood of mating. In our study, there was a strong positive relationship between feeding preference on one or the other hosts and the degree of sexual isolation in the zone of syntopy. Interestingly, the magnitude of divergence was not greater than that observed when individuals from different allopatric populations were tested for willingness to mate. Strong selection for host use in allopatric populations may have increased the amount of divergence because of the absence of alternate hosts and lack of opportunity for gene flow. However, these results are not consistent with the hypothesis that reinforcement is occurring in the zone of syntopy. This conclusion is also supported by the lower strength of divergence in feeding preference in the zone of syntopy. Spatial variation in selection could lead to divergence, the tempo of which depends on a balance between selection and the counteracting effects of gene flow (Slatkin, 1987 ; Nosil, 2005). Gene flow counteracts this divergence between populations with the exception of the situation of reinforcement where gene flow

under limited conditions (is low enough to prevent homogeneity but high enough to promote reinforcement) could enhance selection to avoid hybridization (Nosil, 2005). Thus, gene flow could either enhance or erode divergence. Selection against host switching contributes to divergence in syntopy in Kansas even in the phase of limited gene flow is clear from divergence pattern observed.

Phenotypic divergence in color and cuticular characteristics

Some orthopterans use color for identifying mates where as in others cuticular hydrocarbons have been shown to aid mate recognition (Tregenza and Wedell, 1997) and mediate kin selection (Simmons, 1990). Mullen et al. (2006) found that crickets in the genus *Laupala* exhibit rapid evolution for changes in cuticular lipid composition among species in association with speciation. Assortative mating could result from variation in cuticular hydrocarbons as observed in a hybrid zone between subspecies in the meadow grasshopper *Chorthippus parallelus* (Neems and Butlin, 1994; Tregenza et al. 2000). Mating isolation is key to speciation and results of these studies are indicative of the kinds of roles such components play in eventual reproductive isolation. Results of NIR spectra show that identification of individuals using the near-infrared spectroscopy technique was possible with 96.4 % accuracy in overall comparisons of the two different host plants. Using all the wavelengths in the visual (color) and NIR region gave the most accurate host race identification. But when analyzed separately, classification scores in NIR spectra were lower in the host race that fed on *Gutierrezia* (81.4%) than in the host race that used *Solidago* (98.0%). Nonetheless, our lowest scores were higher than the 76% accuracy observed in similar studies on colony identification of weaver ant

Oecophylla smaragdina using NIR spectroscopy (Newey *et. al*, 2008). The strength of our coefficient of determination (R^2) and classification accuracy were comparable to that of species differences observed between Tobacco budworm and Corn earworm (Jia et al. 2007). Wavelengths of 450-700nm, 900-1400nm and 1500-1700nm have previously been shown to be useful in differentiating insect species (Dowell et al. 1999). Our study identified two additional wavelengths of 1820nm and 1900nm that could be used in differentiating *H. viridis* variants.

Cuticular lipids have been implicated in the mate recognition process of several insect species. (Dapporto, 2007; Tregenza et al. 2000) The differentiation in the NIR spectra observed for individuals feeding on *Gutierrezia* and *Solidago* also suggest the possibility of host based variation in some chemical components of the cuticle in the two insect forms. Hydrocarbons make up major components of cuticular lipids with ~32% of lipids in rice weevils (Baker et al. 1984) and 48-58% in Mormon cricket (Baker et al. 1960). The 1360nm wavelength corresponds to CH3, an important chemical moiety of components that make up the epi-cuticular lipids in insects (Dowell et al. 1999). The 1820nm wavelength corresponds to the structural polysaccharide cellulose which makes up the cell wall in plants (Shenk et al. 1992). Because chitin is the corresponding polysaccharide in insects that makes up the exoskeleton, I hypothesize that 1820nm may represent variation in chitin composition in *H. viridis.* A wavelength of 1900nm corresponds to –Co2H (carboxylic group), a component of the hydrocarbon biosynthesis pathway in insects (Major and Blomquist, 1978). The results of our study reveal significant host based phenotypic differences in color and cuticular composition. Color and cuticular variation observed in the NIR study in the two host forms may aid *H. viridis*

in mate recognition and lead to further divergence of mate preferences. Assortative mating based on host plants could occur through a positive relationship between host choice and mate choice. Alternatively the phenotypic differentiation could also be a by product of host associated selection and divergence.

Conclusions

The observed divergence in feeding and mating preference in association with host use has a strong genetic structure associated with it. Genetic studies conducted in our lab (Grace et al. 2009 ; Chapter 1) using mitochondrial and microsatellite DNA markers revealed a genetic component to the host associated divergence in *H. viridis* populations in Kansas. Mitochondrial DNA results point to a divergence time of ~110,000 years based on molecular clock estimates and isolation in a single refugium followed by radiation and rapid population expansion. Strong selection biased by host plant use was detected for a microsatellite locus, further supporting the notion that host plant use is an important mechanism responsible for structuring *H. viridis* populations in Kansas. No significant variation at neutral loci among the same populations further support the role of divergent selection on host plant use in this system. Maladaptive host switching in the zone of syntopy should result in reinforcement of divergence when compared to allopatry. The lack of reinforcement in the zone of syntopy could be because too little time has passed for reinforcement to occur since the divergence into two host forms. Lack of divergence at neutral loci in genetic studies is consistent with this possibility. Reinforcement could thus be a product of time since isolation, strength of selection and interacting effects of homogenizing gene flow.

Feeding preference tests, mating trials and genetic data (Chapter1) provide evidence for strong selection and adaptation to feed on different host plants as the evolutionary force behind the observed divergence in reproductive isolation. Because patterns of divergence among the allopatric populations did not vary significantly from that of syntopic areas and did not show isolation by distance, it is unlikely that genetic drift played a major role in the observed pattern. Microhabitat selection by *H. viridis* is also highly restrictive with individuals spending most of the time on their respective host plants and only move to short distances to neighboring plants of the same species (Parker, 1982 and 1984 ; Parker and Root 1981). Thus nymphal development and growth, feeding, mating and other activities all take place on an individual's host plants. Such a strong association with the host plant and the associated preference is likely to foster reproductive isolation. Mate choice trials provide strong support that feeding preference promotes reproductive isolation. Shifts to different host plants and subsequent adaptation to use new host species is likely a major factor that contributed to the initial reduction in gene flow between populations. It has also been reported that host preference could evolve in parapatric populations from direct selection on color patterns, through indirect selection (Nosil et al. 2006), providing evidence that phenotypic traits influence selection regimes. I found significant variation among the host forms in color, cuticular composition, feeding preference, mate choice and genetic support for divergence, this provide us evidentiary support for strong selection on phenotypic traits promoting divergence. A host shift enabling feeding divergence and habitat isolation could be the initial barrier to gene flow followed by pre-mating isolation among the two groups with increased sexual isolation. In this study, I was able to identify host associated variants of

H. viridis with high accuracy, suggesting that phenotypic differentiation associated with host plant preferences has occurred. Whether this divergence in phenotypic traits arose as an after effect of increasing mating divergence or has evolved as a result of feeding divergence on alternate hosts is unclear. Nevertheless, phenotypic divergence in color and cuticular characters could further cement mating isolation and potentially promote assortative mating thus enhancing the rate of genetic divergence among the host races. Strong selection could act on all these forces in an additive way to further divergence. Observed host based differences in phenotypic characters, feeding divergence and sexual isolations all suggest populations in an intermediate stage of diversification and speciation. Additional information from genetic data helps us to validate this prediction of diversification. *Hesperotettix viridis* populations thus represent two host-related forms as expected in incipient speciation.

Table 1. A). Chi-squared test results of feeding preference divergence between individuals of Gutierrezia *sp* and Solidago *sp*.
B). Feeding preference divergence among individuals of allopatric vs syntopic populations.
C). Feeding preference divergence among individuals of 8 populations in study zone in Kansas.

	Experiment	***Gutierrezia*** *(Solidago)*	***Solidago*** (*Gutierrezia*)	***Test statistic***	***df***	**P-*value***
A						
1	**Gutierrezia**	163 (6)		145.85	1	<0.0001
2	**Solidago**		110 (20)	64.19	1	<0.0001
B						
3	**All Gut vs All sol**	163 (6)	110 (20)	199.95	1	<0.0001
4	**Allopatric Gut Vs Sol**	63 (3)	57(5)	94.56	1	<0.0001
5	**Syntopy Gut Vs Sol**	100(3)	53(15)	101.31	1	<0.0001
C						
6	**Ellis Gut**	51	1	48.07	1	<0.0001
7	**Russell Gut**	49	2	43.31	1	<0.0001
8	**Hamilton Gut**	28	2	22.53	1	<0.0001
9	**Sherman Gut**	35	1	32.11	1	<0.0001
10	**Ellis Sol**	9	22	5.45	1	0.01
11	**Russell Sol**	6	31	16.89	1	<0.0001
12	**Konza Sol**	2	30	24.5	1	<0.0001
13	**Marysville Sol**	3	27	19.2	1	<0.0001

Insects feeding on the alternate host in paranthesis.

Table 2. Magnitude of sexual isolation (I_{PSI}) among populations of *Hesperotettix viridis.*

	Host Comparison	*I_{PSI} Sexual Isolation Index*	*P value*
1	**RSL Gut vs RSL Sol***	0.78 ± 0.15	0.0004
2	**RSL Gut Vs MYS Sol**	0.84 ± 0.12	0.0002
3	**RSL Sol Vs SHR Gut**	0.75 ± 0.15	0.0006
4	**MYS Sol Vs SHR Gut**	0.89 ± 0.11	0.0002
5	**RSL Gut VS SHR Gut**	-0.040 ± 0.19	0.86
6	**RSL Sol vs MYS Sol**	-0.040 ± 0.18	0.85
7	**RSL Gut Vs RSL Gut**	0.0006 ± 0.18	0.97
8	**RSL Sol Vs RSL Sol**	0.0006 ± 0.18	0.97
9	**SHR Gut Vs SHR Gut**	0.0360 ± 0.18	0.85
10	**MYS Sol Vs MYS Sol**	0.0006 ± 0.18	0.97
11	**ELS Gut Vs ELS Sol***	0.46 ± 0.20	0.03
12	**ELS Gut Vs KNZ Sol**	0.76 ± 0.14	0.0006
13	**ELS Sol Vs HML Gut**	0.68 ± 0.15	0.0008
14	**KNZ Sol Vs HML Gut**	0.84 ± 0.12	0.0002
15	**ELS Gut vs HML Gut**	-0.001 ± 0.18	0.95
16	**ELS SOL vs KNZ Sol**	0.0020 ± 0.18	0.98
17	**ELS Gut Vs ELS Gut**	0.0006 ± 0.18	0.97
18	**ELS Sol Vs ELS Sol**	0.0360 ± 0.18	0.85
19	**HML Gut Vs HML Gut**	0.0006 ± 0.18	0.97
20	**KNZ Sol Vs KNZ Sol**	0.0006 ± 0.18	0.97

Shaded cells : different host mating trials

Table 3. 95% confidence intervals for mating probabilities estimated using Logistic ANOVA procedure in SAS 9.1.3.

	Lower	Mean	Upper
Gutierrezia* vs. *Gutierrezia	0.91	0.97	0.99
Gutierrezia* vs. *Solidago	0.09	0.14	0.21
Solidago* vs. *Solidago	0.93	0.99	0.99

Table 4. Sampling locations and host plant association of *Hesperotettix viridis* populations used in the Near-infrared spectroscopy study.

Population Location (Kansas)	*Host plant*	*Number of Grasshoppers Scanned*		*GPS Co-ordinates*	
				Latitude	Longitude
		Male	Female		
Sherman - ***G1***	*Gutierrezia sp.*	24	25	39.3288	-101.99541
Hamilton - ***G2***		23	28	37.9435	-101.49585
*Ellis - ***G3***		17	27	38.8616	-98.10918
*Russell - ***G4***		25	25	38.8536	-98.525517
Marysville - ***S1***	*Solidago Sp.*	15	25	39.9144	-96.66495
Konza - ***S2***		13	18	39.1013	-96.59833
*Ellis - ***S3***		8	23	38.8616	-98.10918
Russell - ***S4		25	15	38.8536	-98.52551

*Populations in zone of syntopy

Table 5. Host plant prediction using G1 (Sherman) and S1 (Marysville) calibration model to predict G2 (Hamilton), G3(Ellis), G4 (Russell), S2 (Konza), S3 (Ellis), and S4 (Russell).

Population Location	***Host plant***	***Host Plant-Based Prediction Percent Correct Classification***		
		450 -2250 nm	450 -750 nm	750 - 2250 nm
Hamilton (G2)	*Gutierrezia* sp.	98.0	100.0	82.4
Ellis* (G3)	*Gutierrezia* sp.	93.2	95.5	81.8
Russell* (G4)	*Gutierrezia* sp.	94.0	98.0	80.0
Combined G2+G3+G4	*Gutierrezia* sp.	**95.2**	**97.9**	**81.4**
Konza (S2)	*Solidago* sp.	100.0	100.0	96.8
Ellis* (S3)	*Solidago* sp.	100.00	100.0	100.0
Russell* (S4)	*Solidago* sp.	95.0	95.0	97.5
Combined S2+S3+S4	*Solidago* sp.	**98.1**	**98.1**	**98.0**

Table 6. Cross validation and partial least square regression analysis details among *Hesperotettix viridis* populations.

Comparisons/Model	No. of Factors (required for model calibration)	R^2	SECV	Percent Correct Classification		
				Gutierrezia sp. = 1	*Solidago sp.* = 2	Weighted
All Gut vs All Sol G1+G2+ G3+G4 vs S1+S2+S3+S4	8	0.80	0.22	97.9	95.2	96.4
Outside zone of Syntopy Gut vs Sol G1+G2 vs S1+S2	8	0.84	0.20	100	98.6	99.4
Zone of Syntopy Gut vs Sol G3+G4 vs S3+S4	8	0.81	0.21	96.8	98.1	96.7
Gut Vs Gut G1+G2 vs G3+G4	8	0.36	0.40	85.0	78.7	81.9
Sol Vs Sol S1+S2 vs S3+S4	8	0.20	0.45	80.3	64.2	72.6

R^2 : Coefficient of determination ; SECV: Standard error of cross validation

Fig 1 : Proportion of individuals feeding on *Gutierrezia and Solidago* in each population of *Hesperotettix viridis* in Kansas

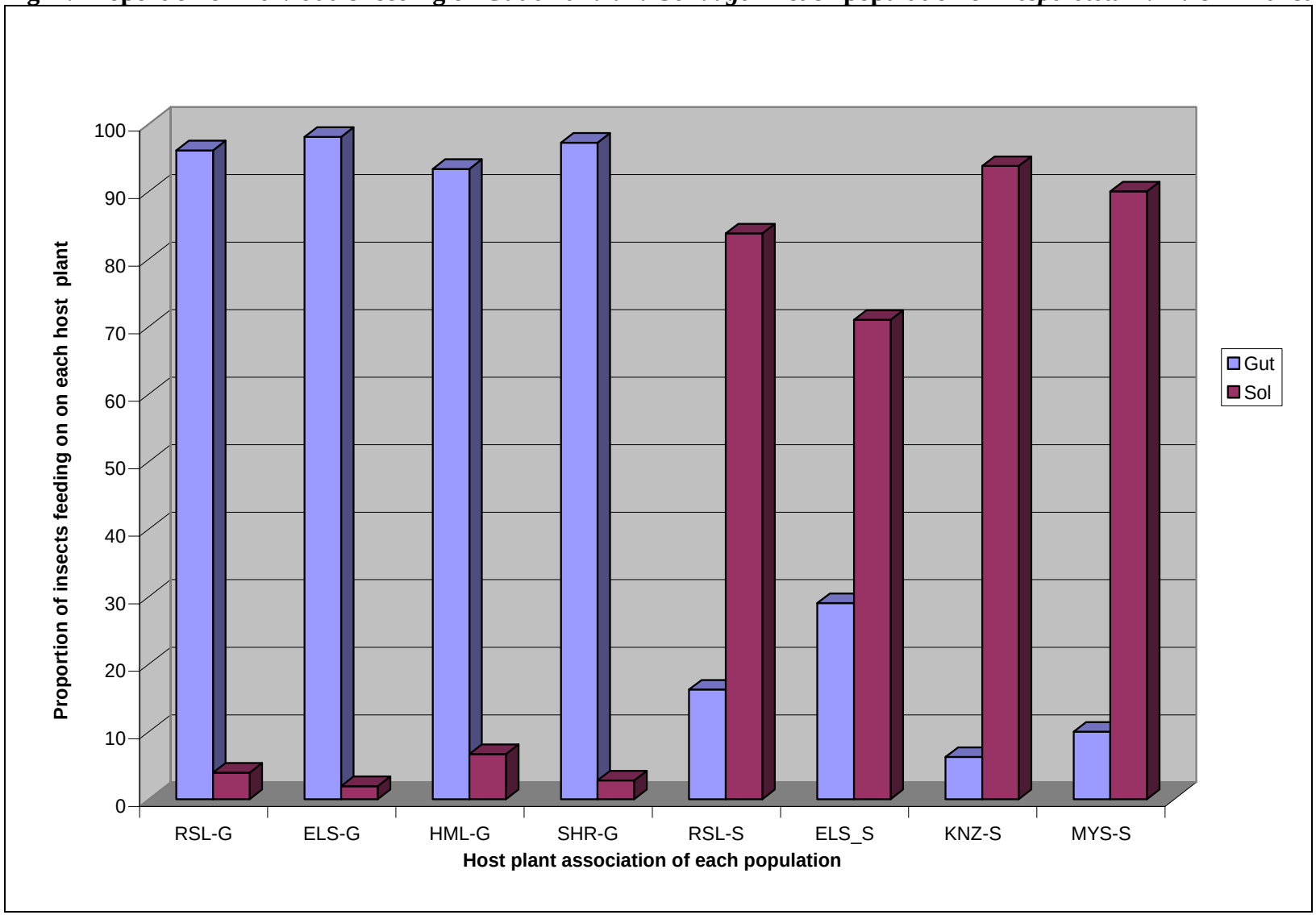


Fig. 2 : Frequency of matings (no-choice trials) in different host and same host trials

G : *Gutierrezia* populations ; S : *Solidago* populations.

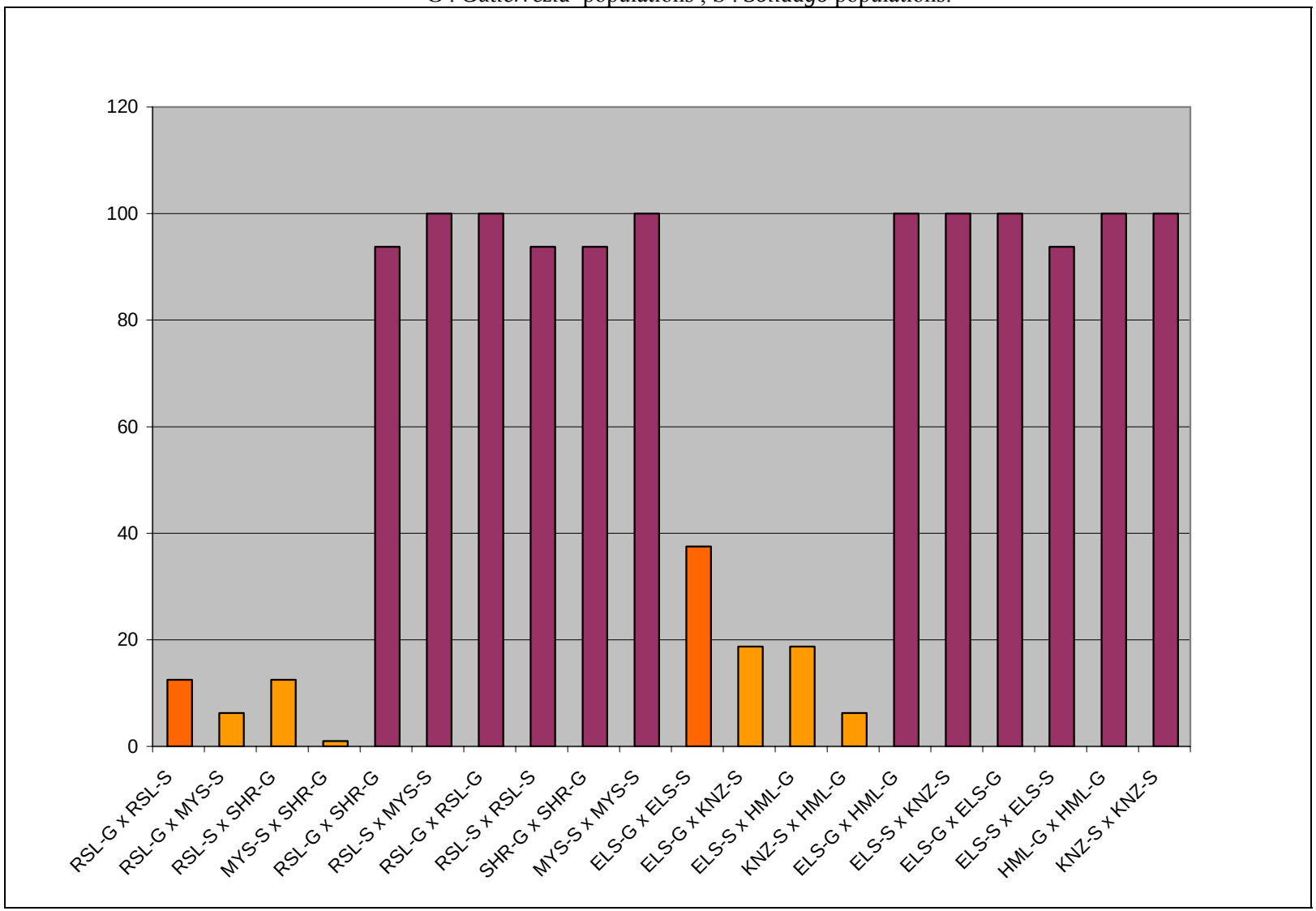
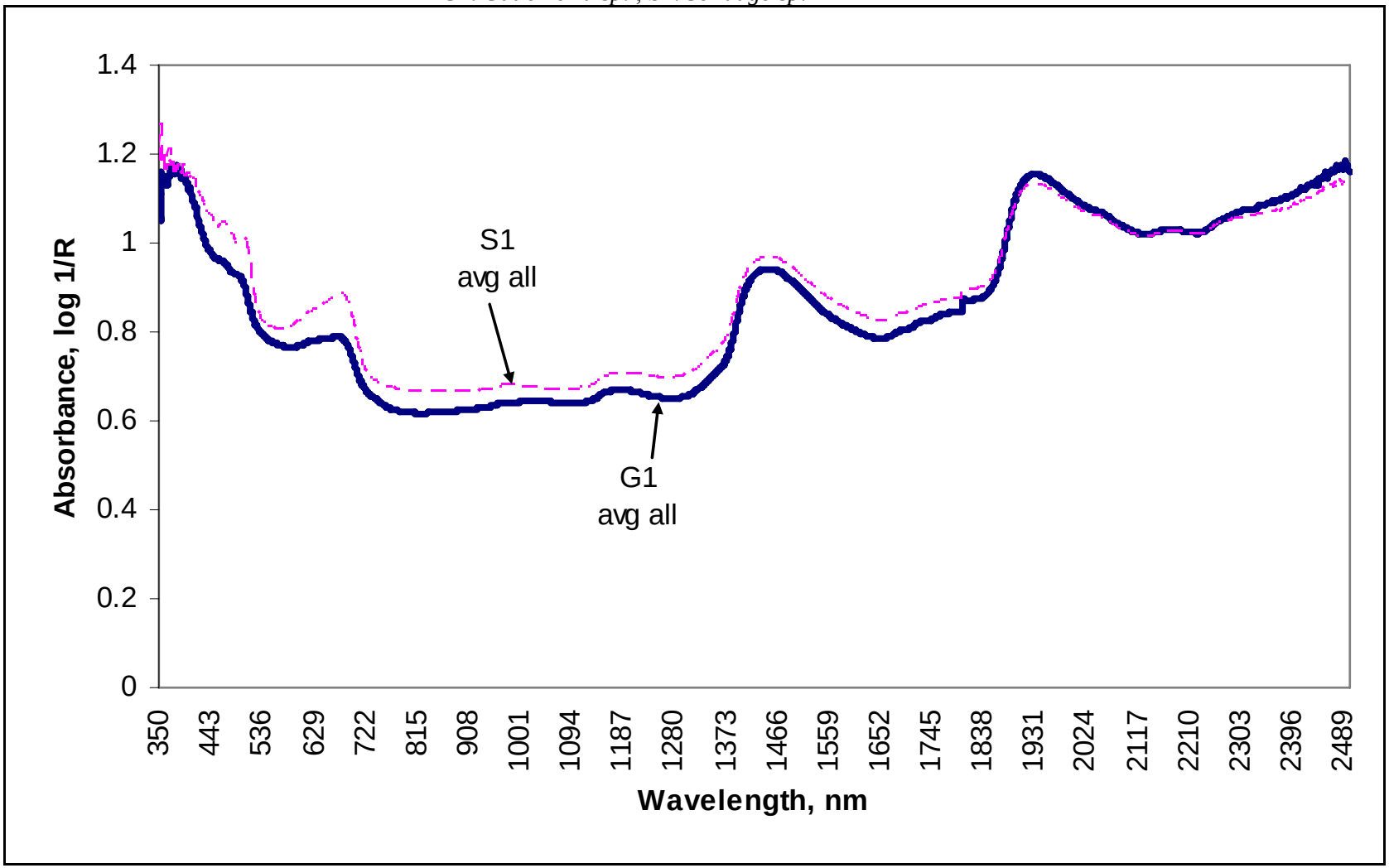


Fig. 3. Average Near-infrared spectra of all *Gutierrezia* and all *Solidago populations* of *H viridis*.

G1: *Gutierrezia sp.* ; S1: *Solidago sp.*



Chapter IV

Research Summary

The current study across a zone of syntopy in Kansas allowed us to assess the extent of host plant association among populations, the extent of sexual isolation and mate discrimination among them, phenotypic divergence among host forms and the underlying genetic divergence among populations to further clarify differentiation patterns associated with host plants in the grasshopper *Hesperotettix viridis*. Analysis of replicated population pairs on different hosts that are ecologically divergent and on same host those are ecologically similar allowed us to determine that host use was the likely cause for the differentiation patterns observed.

Strong feeding preference divergence observed in the two host races were correlated with mate choice with both exhibiting a similar divergence trend. Irrespective of population of origin, insects chose the host plant from which they were collected over the alternate host. This is evidence for habitat isolation driving divergence in feeding preference. Habitat isolation can influence mate choice (Via, 1999) and it is known that this process can be accelerated by divergence in mating signals (Tregenza et al. 2000). As observed with feeding preference, no-choice mating experiments reveal that insects always preferred to mate with the same host plant form rather than with the alternate form. The fact that insects exhibit strong mate choice even in no-choice experiments reveal strong isolating mechanisms or barriers operating in sexual isolation among the two groups. This is further supported by the finding that among the different host pairs that mated, when given a choice, all mated with their own kind. Strong assortative mating based on host plants also suggests that host choice directly influences sexual isolation among the two host forms via habitat isolation. Even though divergence in feeding preference and mate choice in the zone of syntopy were slightly lower than among allopatric populations, these differences were not significantly different among the regions. It is clear from the

magnitude of divergence in these two traits that high sexual isolation and feeding preference divergence occur in *H. viridis.* Color and cuticular variations have been reported to have a role in communication during courtship and thus increase isolation among host races via divergence in these traits. Phenotypic differentiation studies support this possibility in *H. viridis* and reveal strong evidence for divergence in color and cuticular composition among the two host forms. Thus phenotypic divergence may contribute to the process of speciation because divergence of such signals would lead to non-random mating between populations (Neems and Butlin, 1994). In herbivorous insects, diet composition can influence the abundance and composition of hydrocarbons in the cuticle (Chapman et al. 2000). Thus it appears that diet choice has structured the observed phenotypic variations that would act as an additive divergent mechanism furthering sexual isolation among the two host forms.

Sword et al. 2005 and Apple et al., 2009 (submitted) found significant host-based genetic divergence among *H. viridis* populations using AFLP markers and found a role for both genetic drift and selection in the diversification pattern. Our analysis using seven microsatellite markers also found significant host based genetic differentiation and structuring at a selected locus in *H. viridis* populations in Kansas and the results of the study support the role of divergent selection as the evolutionary force behind the observed differentiation pattern. One selected locus revealed clear differences between host plant groups with one allele near fixation in *Solidago* feeding populations and a different allele in higher proportions in *Gutierrezia* populations. This provides strong evidence for divergent selection at that locus. No significant differentiation at the neutral loci either among different host plant groups or same host plant groups suggests that neutral differentiation has nor played a significant role is the divergence pattern. However only a small

section of the genome was sampled which may limit the interpretation of neutral differentiation. However from past studies, and from the results of our study, it is clear than divergent selection is the major evolutionary force structuring host-based diversification in *Hesperotettix viridis.* Gene flow estimates using microsatellites reveal a pattern of progressive expansion of *H. viridis* populations from *Gutierrezia* in the west to *Solidago* populations in the east.

Time since divergence is important in assessing the magnitude of reproductive isolation and genetic divergence existing among the host plant groups. Mitochondrial DNA reveals that divergence is a recent phenomenon with all populations in Kansas originating from a single refugium. There is also evidence for populations within that refugium potentially diversifying and expanding during the rearrangement and expansion of flora and fauna during the late Pleistocene. Irrespective of the fact that less time has passed since divergence, significant differentiation was observed for feeding preference, mate choice and at genetic loci. This is plausible as the higher variance among different host plant groups represents divergence as a result of rapid selection at one (or a few) loci that determines fitness and host-based survival, which translated to the high level of among host-plant group variance. Alternatively genetic drift could structure neutral variation and lead to speciation in herbivorous insects as a result of the highly fragmented distribution of its hosts leading to a fragmented population structure (Mopper et al. 1995). However the fact that neutral loci and mtDNA do not show much variation is possibly a function of the time since divergence into two host forms for drift or mtDNA mutations to structure genetic variation.

An ecological source of divergent selection, a form of reproductive isolation and a genetic mechanism linking these two factors are prerequisites for ecological speciation to occur (Rundle and Nosil, 2005). The results of our study support all these prerequisites. Differences in

habitat and host plant use, as well as sexual selection via divergence in traits involved in mate recognition all serve as ecological sources of divergent selection in this species. Habitat isolation and sexual isolation serve as mating barriers that lead to assortative mating among host plant groups and thus represent mechanisms for reproductive isolation. Genetic divergence at microsatellite loci and AFLP markers (in previous studies) suggest a genetic mechanism that is associated with divergence and potentially connecting the ecological traits to that for reproductive isolation. Progressive expansion of populations from *Gutierrezia* to *Solidago* in Kansas coupled with the short divergence time estimate suggest an ancient host shift occurring within a single refugium followed by rapid population expansion and diversification on two different host plants. Variation in diversification patterns observed in two different pairs of syntopic zone populations in genetic and phenotypic characteristics suggest to varying selection pressures operating at different locations structuring variation in slightly different ways.

Unlike that observed in true specialists, processes that promote genetic divergence are less well studied in oligophagous insect species. This study thus provided a good backdrop of how diversification proceeds in an oligophagous species and the results add several key pieces of information to the overall knowledge base of ecological speciation. Future studies will address if hybridization occurs in the zone of syntopy using mating studies that would test for offspring production and hybrid viability among host plant groups. It will be valuable to undertake a study that addresses genetic mechanisms driving differentiation via a genome scan to identify potential genes involved in diversification. A 454 sequencing approach of cDNA from feeding and mating individuals on both hosts will give comparative evidence for genes activated at these times and thus throw light on the genes involved in host adaptation.

Chapter V

References / Bibliography

Abrahamson, W. G., and Weis, A. E. (1997). Evolutionary ecology across three trophic levels: goldenrods, gallmakers, and natural enemies: Princeton Univ. Press, Princeton, NJ.

Agrawal, A. A. (2001). Phenotypic plasticity in the interactions and evolution of species. *Science.* **294 :** 321 -326.

Arias, C.F., Munoz, A.G., Jiggins, C.D., Mavarez, J., Bermingham, E. and Linares, M. (2008). A hybrid zone provides evidence for incipient ecological speciation in *Heliconis* butterflies. *Molecular Ecology.* **17** : 4699 - 4712.

Arnold, M. L. (1997). Natural hybridization and evolution. Oxford Univ. Press, New York, NY.

Avise, J. C. (1994). Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Avise, J. C., Walker, D. and Johns, G. C. (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London. B Biol.* **265** : 1707-1712.

Baker, G. L., Pepper, J. H., Johnson, L. H. and Hastings, J. E. (1960). Estimation of the composition of the cuticular wax of the Mormon cricket, *Anabrus simplex* Hald. *Journal of Insect Physiology*. **5** : 47-60.

Baker, J. E., Woo, S. M., Nelson, D. R. and Fatland, C. L. (1984). Olefins as major components of epicuticular lipids of three *Sitophilus* weevils. *Comparitive. Biochemistry and Physiology.* B. **77**: 877-884.

Bandelt, H.J., Forster, P. and Röhl, A. (1999). Median–joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution.* **16 :** 37–48.

Barton, N. H. and Gale, K. S. (1993). Genetic analysis of hybrid zones. In *Hybrid zones and the evolutionary process*. Oxford Univ. Press, Oxford, U.K. R. G. Harrison, ed. pp. 13-45.

Barton, N. H. and Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology Evolution & Systematics.* **16**:113-148.

Beerli, P. and Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America.* **98 :** 4563–4568.

Bensasson, D., Zhang, D.X. and Hewitt, G.M. (2000). Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Molecular Biology and Evolution.* **17:** 406–415.

Berenbaum, R.M. and Zangerl, A.R. (1998). Population level adaptation to host plant chemicals: The role of cytochrome P450 monoxygenases. In- *Genetic Structure and Local*

Adaptation in Natural Insect Populations: Effects of Ecology, Life History and Behavior. Chapman & Hall, London. Mopper, S & Strauss, S.Y. eds.

Berlocher, S. and Feder, J. (2002). Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology*. **47**: 773-816.

Bernays, E.A and Chapman, R.F. (1994). *Host Plant Selection by Phytophagous Insects.* Chapman & Hall, NY.

Boecklen, W.J and Mopper, S. (1998). Local adaptation in specialist herbivores: Theory and Evidence. In- *Genetic Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology, Life History and Behavior*. Chapman & Hall, London. Mopper, S. & Strauss, S.Y. eds.

Bohonak, A. J. 2002. IBD (Isolation By Distance): A program for analyses of isolation by distance. *Journal of Heredity.* **93**: 153-154.

Bonin, A., Taberlet, P., Miaud, C. and Pompanon, F. (2006). Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution.* **23**:773–783.

Brandon, R.N. (1990). Adaptation and environment. Princeton Univ. Press. Princeton, NJ. 214 pp.

Brower, A.V.Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Science U S A.* **91**: 6491–6495.

Brown, W. L. and Wilson, E. O. (1956). Character displacement. *Systematic Zoology.* **5** : 49-65.

Burnham, K. P. and Anderson, D. R. (1998). Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, USA. 353 pp.

Bush, G. L. (1994). Sympatric speciation in animals: New wine in old bottles. *Tree.* **9** : 285–288.

Butlin, R., Bridle, J.R. and Schulter, D. (2009). *Speciation and Patterns of Diversity.* Cambridge University Press.

Carstens, B.C. and Knowles, L.L. (2006). Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology*, **16**, 619-627.

Chapman, R. F., Espelie, K. E. and Peck, S.B. (2000). Cuticular hydrocarbons of grasshoppers from the Galapagos Islands, Ecuador. *Biochemical Systematics & Ecology*. **28** (6) : 579-588.

Chapman, R.F., Espelie, K.E and Sword, G.A. (1995). Use of Cuticular lipids in grasshopper taxonomy: A study of variation in *Schistocerca Shoshone* (Thomas). *Biochemical Systematics & Ecology*. **23** (4) : 383-398.

Coyne, J. A and Orr, H.A. (2004). *Speciation.* Sinauer Associates Sunderland, MA.

Dallas, J. F. (1992). Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome.* **3** : 452-456.

Dapporto, L. (2007). Cuticular lipid diversification in *Lasiommata megera* and *Lasiommata paramegaera*: the influence of species, sex, and population (Lepidoptera: Nymphalidae). *Biological Journal of the Linnaean Society.* **91**:703–710.

Dieckmann, U. and Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*. **400** : 354–357.

Dietrich, W., Katz, H., Lincoln, S.E., Shin, H.S., Friedman, J., Dracopoli, N.C. and Lander, E.S. (1992). A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**: 423–447.

Dobzhansky, T. (1951). *Genetics and the Origin of Species.* Columbia Univ. Press. New York. 3rd ed

Doebeli, M. (1996). A quantitative genetic competition model for sympatric speciation. *Journal of Evolutionary Biology*. **9** : 893–909.

Doebeli, M., and Dieckmann, U. (2000). Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *American Naturalist*. **156** : S77–S101.

Dowell, F. E., Throne, J. E., Wang, D and Baker, J. E. (1999). Identifying stored-grain insects using Near-infrared spectroscopy. *Journal of Economic Entomology.* **92**: 165-169.

Dre`s, M. and Mallet, J. (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the* Royal *Society of London B.* **357** : 471-492.

Egan, S.P., Nosil, P. and Funk, D.J. (2008) Selection and genomic differentiation during ecological speciation: isolating the contributions of host-association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. *Evolution* , **62,** 1162-1181, 2008

Ellengren, H. (1995). Mutation rates at porcine microsateliite loci. *Mammalian Genome*. **6**. 376-377.

Eriksson, G. (1998). Evolutionary forces influencing variation among populations of *Pinus sylvestris*. *Silva Fennica*. **32** (2) : 173-184.

Excoffier, L. and Slatkin, M. (1998). Incorporating genotypes of relatives into a test of linkage disequilibrium. *American Journal of Human Genetics.* **62**: 171-180.

Excoffier, L., Laval, G. and Schneider, S. (2005). Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1:** 47–50.

Feder, J. L. (1998). The apple maggot fly, *Rhagoletis pomonella*: Flies in the face of conventional wisdom about speciation? *In* D. J. Howard and S. H. Berlocher (eds.), *Endless forms: Species and speciation*, pp. 130–144. Oxford University Press, Oxford.

Feder, J. L. and Forbes, A. A. (2007). Habitat avoidance and speciation for phytophagous insect specialists. *Functional Ecology*. **21** : 585 -597.

Feder, J. L., Berlocher, S. J., Roethele, J. B., Dambroski, H., Smith, J. J., Perry, W. L., Gavrilovic, V., Filchak, K. E., Rull, J. and Aluja, M. (2003a). Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences*. **100**: 10314–10319.

Feder, J. L., Roethele, F. B., Filchak, K., Niedbalski, J. and Romero-Severson, J. (2003). Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics.* **163** : 939–953.

Feder, J. L., Xie, X., Rull, J., Velez, S., Forbes, A., Dambroski, H., Filchak, K. and Aluja, M. (2005). Mayr, Dobzhansky, Bush and the complexities of sympatric speciation in Rhagoletis. *Proceedings of the National Academy of Sciences*. **102**:6573– 6580.

Feder, J. L. and Forbes, A. A. (2008). Host fruit odor discrimination and sympatric host race formation. *In Evolutionary biology of Herbivorous insects, Specialization, Speciation and Radiation* University of California Press. pp 341. Tilman, K.J. ed.

Felsenstein, J. (2005). PHYLIP. http://evolution.genetics.washington. edu/ phylip.html.

Ferreira, M., Karunaweera. N., da Silva-Nunes, M., da Silva, N.S., Wirth, D. and Hartl, D. (2007). Population structure and transmission dynamics of *Plasmodium vivax* in rural Amazonia. *Journal of Infectious Diseases.* **195** : 1218–1226.

Fox, L.R. and Morrow, P.A. (1981). Specialization: species property or local phenomenon? *Science.* **211 :** 887–893.

Funk, D. J and Nosil, P. (2008). Comparitive analysis of Ecological speciation. *In Evolutionary biology of Herbivorous insects, Specialization, Speciation and Radiation.* University of California Press. Pp 341. Tilman, K.J. ed.

Funk, D. J., Nosil, P. and Etges, W.J. (2006). Ecological divergence is consistently positively associated with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences,* USA. **103:** 3209-3213.

Funk, D.J. (1998). Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution.* **52** :1744–1759.

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

Galactic Industries. (1996). Grams/32 user's guide, version 4.0.Galactic, Salem, NH.

Gavrilets, S. and Vose, A. (2009). Dyanamic patterns of adaptive radiation: evolution of mating preferences. In *Speciation and Patterns of Diversity*. Cambridge University Press. Butlin, R; Bridle, J.R ; Schulter, D. eds.

Geritz, S. A., Kisdi, H. É., Meszéna, G and Metz, J. A. J. (1999). Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology.* **12** : 35–57.

Gorur, G., Lomonaco, C. & Mackenzie, A. (2005). Phenotypic plasticity in host plant specialization in *Aphis fabae*. *Ecological Entomology.* **30 :** 657–664.

Gorur, G., Lomonaco, C. and Mackenzie, A. (2007). Phenotypic plasticity in host choice behavior in black bean aphid, *Aphis fabae* (Homoptera: Aphididae). *Arthropod-Plant Interactions.* **1** (3) : 187-194.

Goudet, J. (2001). F_{STAT}, a program to estimate and test gene diversities and fixation indices (version 2.9.3).

Grace, T., Joern, A., Apple, J., Brown, S.J. and Wisely, S. (2009). Highely polymorphic microsatellites in the North American Snake weed Grasshopper, *Hesperotettix viridis*. *Journal of Orthoptera Research* **18**(**1**) In press.

Hamilton, M.B., Pincus, A., Di Fiore, E.L. and Flesher, R.C. (1999). Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques*. **27 :** 500–507.

Haverty, M.I., Page, M., Nelson, L.J and Blomquist, G.J. (1988). Cuticular hydrocarbons of dampwood termites, *Zootermopsis*: intra- and intercolony variation and potential as taxonomic characters. *Journal of Chemical Ecology* **14** : 1035-1058.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological. Journal of the Linnaean. Society*. **58** : 247-276.

Hewitt, G. M. (2000). The genetic legacy of the quaternary Ice Ages. *Nature.* **405** : 907-913.

Itami, J.K., Craig,T.P. and Horner, J.D. (1998). Factors affecting gene flow between the host races of *Eurosta solidaginiss*. In- *Genetic Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology, Life History and Behavior*. Chapman & Hall, London. Mopper, S. & Strauss, S.Y. eds.

Jia, F., Maghirang, E.B., Dowell, F.E., Abel, C and Ramaswamy, S. (2007) Differentiating Tobacco Budworm and Corn Ear worm using Near- Infrared Spectroscopy. *Journal of Economic Entomology* **100** (**3**) : 759-764.

Jiggins, C.D. and Mallet, J. (2000). Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, **15** : 250–255.

Joern, A. (1979). Feeding patterns in grasshoppers (Orthoptera: Acrididae): factors influencing diet specialization. *Oecologia* **38** : 325-347.

Joern, A. (1985). Grasshopper dietary (Orthoptera: Acrididae) from a Nebraska Sand Hills prairie. – *Transactions of the Nebraska Academy of Sciences*. **13**: 21-32.

Joshi, A and Thompson, J.N. (1995). Tradeoffs and the evolution of host specialization. *Evolutionary Ecology*, **9 :** 82–92.

Kimura, M. (1996). Quaternary paleogeography of the Ryukyu Arc. *Journal of Geography.* **105** : 259-285.

Knowles, L. L. and Richards, C.L. (2005). Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Molecular Ecology.* **14 :** 4023-4032.

Larsson, S. and Ekbom, B. (1995). Oviposition mistakes in herbivorous insects: Confusion or a step towards a new host plant? *Oikos.* **72** :155–160.

Lessa, E. P., Cook, J. A and Patoon, J. L (2003). Genetic foot prints of demographic expansion in North America, but not Amazonia during the late Quaternary. *Proceedings of the National Academy of Sciences.* **100**(18) : 10331-10334.

Lesica, P. and Allendorf, F. W. (1995). When are peripheral populations valuable for conservation?. *Conservation Biology*. **9**: 753–760.

Lewis, P.O., and Zaykin, D. (1999). Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0

Lockey, K.H. and Metcalfe, N.B. (1998). Cuticular hydrocarbons of adult *Himatismus* species and a comparison with 21 other species of adult Tenebrionid beetle using multivariate analysis. *Compartive Biochemistry and Physiology.* **91B** : 371-382.

Machado, R.D., Povoa. M., Calvosa, V. P., Ferreira, M., Rossit, A.B., dos Santos, E. M, and Conway, D. (2004). Genetic structure of *Plasmodium falciparum* populations in the Brazilian Amazon region. *Journal of Infectious Diseases*. **190** : 1547–1555.

Major, M. A. and Blomquist, G. J. (1978). Biosynthesis of hydrocarbons in insects: Decarboxylation of long chain acids to *n* –alkanes in *Periplaneta. Lipids.* **13** (5): 323-328.

Marshall, J. L., Arnold, M. L and Howard, D. J. (2002). Reinforcement the road not taken. *Trends in Ecology and Evolution.* (**17**) : 12-558-563.

Michel, A. P., Rull, J., Aluja, M. and Feder, J.L. (2007). The genetic structure of hawthorn-infesting *Rhagoletis pomonella* populations in Mexico: implications for sympatric host race formation. *Molecular Ecology.* **16**: 2867-2878.

Mopper, S. (1996). Adaptive Genetic Structure in Phytophagous insect populations. *Trends in Ecology and Evolution.* **11 :** 235-238.

Mopper, S and Strauss S,Y. eds (1998*) Genetic Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology, Life History and Behavior.* Chapman & Hall, London.

Mullen, S.P., Tamra, C., M., Schal, C and Shaw, K.L (2006). Rapid Evolution of Cuticular Hydrocarbons in a Species Radiation of acoustically diverse Hawaiian Crickets (GRYLLIDAE: TRIGONIDIINAE: *LAUPALA*). *Evolution.* **61**(1) : 223-231.

Muller, H.J. (1942). Isolating mechanisms, evolution and temperature. *Biological Symposium*. **6** : 71-125.

Neems, R. M., and Butlin, R. K. (1994). Variation in cuticular hydrocarbons across a hybrid zone in the grasshopper *Chorthippus parallelus*. *Proceedings of the Royal Society of London*. B. **257** (1349) :135–149.

Newey, P.S., Robson, S.K.A and Crozier, R.H. (2008) Near-infrared spectroscopy identifies the colony and nest of origin of weaver ants, *Oecophylla smaragdina. Insectes Sociaux.* **55** : 171–175.

Noor, M.A.F. (1995). Speciation driven by natural selection. *Nature*, **375** : 674-675.

Noor, M.A.F. (1999). Reinforcement and other consequences of sympatry. *Heredity.* **83 :** 503-508.

Nosil, P. (2005). The role of selection and gene flow in the evolution of sexual isolation in *Timema* walking sticks and other Orthopteroids. *Journal of Orthoptera Research.* **14 :** 247 - 253.

Nosil, P. (2007). Synthesis: divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. *American Naturalist.* 169 : 151-162.

Nosil, P.**,** Egan, S. P. and Funk, D. J. (2008). Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation by adaptation' and multiple roles for divergent selection. *Evolution* **62** : 316-336.

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

Nosil, P., Crespi, B. J. and Sandoval, C. (2002). Host plant adaptation drives parallel evolution of reproductive isolation. *Nature.* **417** : 441-443.

Orcutt, D. M., Nilson, E. T. and Hale, M. G. eds (2000). Herbivory and plant stress ; *Physiology of plants under stress: Soil and Biotic factors*. John Wiley and Sons. 696pp.

Ortiz-Barrientos, D., Counterman, B.A., and Noor, M. A. F. (2004). The genetic basis of speciation by reinforcement. *PLOS Biology*: 2256-2263.

Otte, D. and Joern, A. (1977). On feeding patterns in desert grasshoppers and the evolution of specialized diets. – *Proceedings of the Academy of Natural Sciences.* Philadelphia **128** : 89-126.

Parker, M. A. (1983). Foraging behavior of a specialist grasshopper, *Hesperotettix niridis,* and its impact on the population dynamics of an arid grassland shrub. *Gutierrezia microcephala* (Compositae). Dissertation. Cornell Univ., Ithaca, NY.

Parker, M. A. (1982). Association with mature plants protects seedlings from predation in an arid grassland shrub, *Gutierrezia microcephala*. *Oecologia.* 53 : 276–280.

Parker, M. A. (1984). Local food depletion and the foraging behavior of a specialist grasshopper *Hesperotettix viridis*. *Ecology.* 65: 824–835.

Parker, M. A. and Root, R.B. (1981). Insect herbivores limit habitat distribution of a native composite, *Machaeranthera canescens*. *Ecology.* 62 : 1390–1392.

Payseur, B. A., Cutter, A. D. and Nachman, M. W. (2002). Searching for evidence of positive selection in the human genome using patterns of microsatellite variability. *Molecular Biology and Evolution.* **19** : 1143– 1153.

Peakall, R., Smouse, P.E., (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. **6 :** 288-295.

Peterson, M.A. and Denno, R.F. (1998). Life history strategies and the genetic structure of phytophagous insect populations. Pp. 263-322 in: *Genetic structure and local adaptation*

in natural insect populations: Effects of ecology, life history, and behavior. Mopper, S & Strauss,S eds). Chapman & Hall, New York, NY.

Pfadt, R. E. (1994). Field guide to common western grasshoppers, 2nd ed. Wyoming Agriculture Experiment Station, Laramie, WY.

Phillips, A., Sbatini, A., Milligan, P.J.M., Broomfield, G and Molyneux, D.H. (1990). The *Anopheles maculipennis* complex (Diptera: Culicidae) comparison of the cuticular hydrocarbon profiles determined in adults of five Palaearctic species. *Bulletin of Entomological. Research.* **80**: 459-464.

Piry, S., Luikart, G., and Cornuet, J.M. (1999). BOTTLENECK ; a computer program for detecting recent introductions in the effective population size using allele frequency data. *Journal of Heredity.* **90 :** 502-503.

Pritchard, J, K., Stephens, M. and Donnelly, P. (2000a). Inference of population structure using multilocus genotype data. *Genetics.* **155**, 945–959.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.

Rolán - Alvarez, E & Caballero, A. (2000). Estimating sexual selection and sexual isolation effects from mating frequencies. *Evolution* **54** (1): 30-36.

Runck, A. N and Cook, J. A. (2005). Post-glacial expansion of the southern red-backed vole (*Clethrionomys gapperi*) in North America. *Molecular Ecology.* **14** (5) : 1445–1456.

Rundle, H. and Nosil, P. (2005). Ecological speciation. *Ecology Letters.* **8** : 336-352.

SAS Institute (1990). *SAS Language: References. Version*. SAS: Cary, NC.

Scheffer, S. J. and Hawthorne, D. J. (2007). Molecular evidence of host-associated genetic divergence in the holly leafminer *Phytomyza glabricola* (diptera: agromyzidae): apparent discordance among marker systems. *Molecular Ecology.* **16** : 2627-2637.

Schuelke, M. (2000). An ergonomic method for the fluorescent labeling of PCR fragments, A poor mans approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology*. **18**: 233-234.

Schug, M. D., Hutter, C. M., Wetterstrand, K. A., Gaudette, M. S., Mackay, T.F.C. and Aquadro, C. F. (1998). The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. *Molecular Biology and Evolution.* **15** :1751–1760.

Sequeira, F., Alexandrino, J., Rocha, S., Arntzen, J. W. and Ferrand, N. (2005). Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology.* **14** : 245-254.

Servedio, M. R and Noor, M. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics.* **34** : 339-364.

Servedio, M. R. (2005). The What and Why of research on reinforcement. *PLoS Biology.* **2** : 2032- 2035.

Shenk J. S., Workman, J. J, Jr. and Westerhaus, M. O. (1992). Application of NIR spectroscopy to agricultural products. In: Burns, D.A, and Ciurczak, E.M. editors. *Handbook of Near-infrared analysis*, pp. 383–431. Marcel Dekker.

Simmons, L. W. (1990). Pheromonal cues for the recognition of kin by female field crickets, *Gryllus bimaculatus*. *Animal. Behavior*. **40** :192–195.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*. **236** : 787–792.

Smadja, C and Butlin, R. (2006). The new role for reinforcement. *Heredity* **96 :** 422-423.

Stireman, J. O., Nason, J. D. and Heard, S.B. (2005). Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod-insect community. *Evolution*: **59** (12) : 2573–2587.

Sword, G. A., Joern, A. and Senior, L.B. (2005). Host plant-associated genetic differentiation in the snakeweed grasshopper, *Hesperotettix viridis* (Orthoptera: Acrididae). *Molecular Ecology.* **14** (7) : 2197–2205.

Sword, G. A., Senior, L. B., Gaskin, J. E. and Joern, A**.** (2007). Double trouble for grasshopper molecular systematics: Intra-individual heterogeneity of both mitochondrial 12S-valine-16S and nuclear internal transcribed spacer ribosomal DNA sequences in *Hesperotettix viridis* (Orthoptera: Acrididae). *Systematic Entomology.* 32 : 420-428.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. **123 :** 585–595.

Tamura, K. (1992). The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. *Molecular Biology and Evolution.* **9** : 814–825.

Templeton, A. R. (1980). The theory of speciation via the founder principle. *Genetics* **94**:1011–1038.

Thomas, C. D. and Singer, M. E. (1998). Scale-dependent evolution of specialization in a checker spot butterfly: from individuals to metapopulations and ecotypes. In: Mop per, S.

and Strauss, S. Y. (eds), *Genetic structure and local adaptation in natural insect populations.* Chapman and Hall. New York. pp, 343-374.

Traxler, M.A, and Joern, A. (1999). Performance tradeoffs for two hosts within and between populations of the oligophagous grasshopper *Hesperotettix viridis* (Acrididae). *Oikos.* **87 :** 239–250.

Tregenza, T., and Wedell, N (1997). Definitive evidence for cuticular pheromones in a cricket. *Animal Behavior* **54** : 979–984.

Tregenza, T., Pritchard, V. L. and Butlin, R. K. (2000). The origins of premating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. *Evolution.* **54 :** 1687–1698.

Trontelj, P., Machino, Y. and Sket, B. (2005). Phylogenetic and Phyleogeographic relationships in the cray fish genus Austropotamobium inferred from mitochondrial CO1 gene sequences. *Molecular Phylogenetics and Evolution.* **34 :**. 212-226.

Venkatesan, M., Westbrook, C. J., Hauer, M. C. and Rasgon, J. L. (2007). Evidence for a population expansion in the West Nile virus vector *Culex tarsalis*. *Molecular Biology & Evolution.* **24 :** 1208–1218

Via, S. (1999). Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution.* **53** : 1446–1457.

Vigouroux, Y., McMullen, M., Hittinger, C. T., Houchins, K. and Schultz, L. (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proceeding of the National Academy of Sciences.* **99 :** 9650-9655.

Vitalis, R., Dawson, K., Boursot, P. and Belkhir, K. (2003). DetSel 1.0: A Computer Program to Detect Markers Responding to Selection. *Journal of Heredity*. **94**: 429-431.

Wake, D. B. (1997). Incipient species formation in salamanders of the *Ensatina* complex. *Proceeding of the National Academy of Sciences*.USA **94** :7761-7767.

Weir, B. S. and Cockerham, C. C.. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution.* **38** :1358–1370.

Weiss, S., Knizhin, I., Kirilov, A. and Froufe, E. (2006). Phenotypic and genetic differentiation of two major phylo-geographical lineages of arctic grayling *Thymallus arcticus* in the Lena River, and surrounding Arctic drainages. *Biological Journal of the Linnaean. Society.* **88 :** 511–525.

Wright, S. (1951). The genetical structure of populations. *Annals. Eugenics* **15** : 323–354.

Zhang, D.X. and Hewitt, G.M. (1996). Nuclear integrations: challenge for mitochondrial DNA markers. *TREE.* **11**(6): 10031-10038.

Appendix A: Supplementary Table

Table A-1. Pair wise genetic differentiation (F_{ST}) values among 14 populations of *Hesperotettix viridis* in Kansas estimated from microsatellite data

	Ellis Gut.	**Ellis** Sol.	**Ellsw.** Sol.	**Cimm.** Gut.	**Sew.** Gut.	**Hamil.** Gut.	**Sher.** Gut.	**Hoxie** Gut.	**Ness** Sol.	**Rus.** Gut.	**Rus.** Sol.	**Konz.** Sol.	**Marr.** Sol.	**Mary.** Sol.
Ellis Gut.	*	0.04	0.05	0.06	0.02	0.03	0.03	0.04	0.05	0.05	0.05	0.04	0.03	0.04
Ellis Sol.	0.44	*	0.02	0.04	0.02	0.03	0.03	0.02	0.02	0.04	0.03	0.02	0.02	0.03
Ellsworth Sol.	0.60	0.05	*	0.04	0.02	0.03	0.03	0.03	0.02	0.04	0.03	0.02	0.03	0.03
Cimarron Gut.	-0.03	0.59	0.82	*	0.02	0.02	0.02	0.04	0.04	0.03	0.05	0.03	0.05	0.04
Seward Gut.	0.00	0.59	0.75	-0.04	*	0.01	0.01	0.02	0.02	0.02	0.03	0.02	0.02	0.02
Hamilton Gut.	0.03	0.58	0.70	-0.01	-0.01	*	0.01	0.02	0.03	0.03	0.03	0.02	0.02	0.02
Sherman Gut.	0.08	0.68	0.79	0.03	0.01	0.00	*	0.01	0.02	0.03	0.03	0.01	0.01	0.03
Hoxie Gut.	-0.01	0.63	0.81	-0.03	-0.02	-0.02	0.01	*	0.03	0.04	0.04	0.01	0.01	0.02
Ness Sol.	0.42	-0.01	0.07	0.55	0.56	0.56	0.66	0.59	*	0.04	0.03	0.02	0.03	0.03
Russell Gut.	0.05	0.64	0.77	0.01	0.00	-0.01	-0.01	-0.02	0.62	*	0.04	0.04	0.04	0.04
Russell Sol.	0.51	0.00	0.01	0.71	0.67	0.64	0.74	0.72	0.01	0.70	*	0.02	0.03	0.02
Konza Sol.	0.57	0.00	0.01	0.74	0.71	0.68	0.77	0.76	0.01	0.74	-0.01	*	0.01	0.01
Marrion Sol.	0.47	0.01	0.00	0.66	0.63	0.61	0.71	0.67	0.02	0.67	-0.02	-0.01	*	0.03
Marysville Sol.	0.44	-0.01	0.06	0.57	0.58	0.58	0.67	0.61	-0.01	0.64	0.02	0.02	0.02	*

Above diagonal : Pair wise F_{ST} values Nuetral Loci Below diagonal : Pair wise F_{ST} values Selected Locus (Arlequin v 3.1)