Spring 2023

Genetic Mechanisms Underlying Maladaptation in Specialized Species Interactions

Nitin Ravikanthachari

GENETIC MECHANISMS UNDERLYING MALADAPTATION IN SPECIALIZED SPECIES INTERACTIONS

by

Nitin Ravikanthachari

Bachelor of Science
St. Joseph's College (Autonomous), 2014

Master of Science
Bangalore University, 2016

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Biological Sciences

College of Arts and Sciences

University of South Carolina

2023

Accepted by:

Carol L Boggs, Major Professor

Joseph M Quattro, Committee Member

Ward B Watt, Committee Member

Brian D Hollis, Committee Member

Peter Andolfatto, Committee Member

Cheryl L. Addy, Interim Vice Provost and Dean of the Graduate School

DEDICATION

To Thomas Reid Davys Bell (1863-1948), an exemplary naturalist whose work on early stages of Indian butterflies (1909-1927) influenced my passion for natural history and inspired me to study plant-insect interactions.

ACKNOWLEDGEMENTS

I owe my deepest and heartfelt thanks to Carol Boggs. My growth in the last 5 years would not have been possible without her guidance and encouragement. She has been my biggest supporter and my greatest advocate. Her belief in me and my abilities has been my biggest source of encouragement in this journey. She was always patient, understanding and allowed the utmost independence to pursue my scientific ideas. She always encouraged me to dream big, pursue ideas that seemed difficult and put her faith in me to accomplish them and nudged me in the right directions when I was lost. She also helped me develop my mentoring skills by encouraging me to work with various undergraduate students in building their project ideas and scientific skills. I will forever be grateful to Carol for being the best mentor I could possibly hope for and being my greatest ally in academia.

I thank my committee for all their help and support through my PhD. They provided feedback, encouraged me to pursue new ideas and were patient with their time when I needed it the most. My work has greatly benefitted from their advice and feedback.

My dissertation would not have been possible without the help from all the field assistants and undergraduate researchers who played a crucial role in data collection and field work. I am fortunate to have worked with such great scientists and will forever cherish the journey and experiences with them. Thank you, Malia

Olson, Hannah Walton, Emma Wagner, Skyler McDaniel, Ojaswee Shrestha, Kelley Asselin, Nimue Shive, Daria Syskine, Andrea Chau, Anna Michelle Jones, and Laura Littleton, for everything you have done to help me with my field work and data collection.

My time at UofSC has been memorable and filled with wonderful experiences thanks to the great colleagues in the Biology Department, GABS, and the Biology Administration. Your support has meant a lot to me, and I will always be thankful for that. I also thank my office mates Jesus M Rivera-Hernandez, Chloe Keck, Blair Stokes, Laurent Duverglas, Ojaswee Shrestha, and Rachel Steward for their companionship and for their help both in and out of academia. I am grateful to have shared my time with you. Thanks also to everyone in the Boggs and Watt lab for all the great memories and help in the last five years.

My dissertation involved a lot of sequencing as well as the use of high-performance clusters to analyze the data. I thank the UofSC Research Computing Infrastructure for all their help with my bioinformatic needs. Nathan Elgar and Myk Milligan were particularly helpful in troubleshooting all my bioinformatic needs.

My development through the course of my PhD was enriched by the various discussions and advice that I have received. Carrie Wessinger, Dan Speiser, Nate Senner, Maria Stager, and Meredith Blackwell have been generous with their time and have entertained my ideas and guided me in the right direction.

I also thank my friends at the Rocky Mountain Biological Laboratory (RMBL) who made my fieldwork enjoyable. Thank you Sam Pierce, Justin Bain, Nick Alexandre, Annika Nelson, Jackie Fitzgerald, Elsa Godtfredsen and billy barr for the providing me with a lifetime of memories from my fieldwork.

Finally, I thank my wife Minoti Asher, who has been instrumental in my success during my PhD. I thank her for supporting me and joining me on all my adventures and being my biggest cheerleader. I also thank my parents and sisters who have been the backbone of my support and encouragement.

ABSTRACT

Adaptation has been used as the framework to understand the power of natural selection in structuring biodiversity and driving biological interactions. However, recent examples have indicated that many organisms are maladapted to their local conditions. Maladaptation can result from various causes including rapid environmental change, novel environments, interplay of evolutionary forces (*viz*: gene flow-selection balance) and anthropogenic disturbance.

To dissect the genetic mechanisms underlying maladaptation, I used the interaction between *Pieris macdunnoughii* (Remington) (Lepidoptera: Pieridae) which lays eggs on an invasive mustard *Thlaspi arvense* (L.) (Brassicaceae), which is lethal to the larvae.

First, local adaptation to avoid laying eggs on or adapt to feed on *T. arvense* could be dampened by maladaptive gene flow. To evaluate this, I used a population genomics approach to identify population structure, quantify gene flow, and identify signatures of local adaptation. I found signatures of local adaptation in the face of rampant gene flow highlighting the potential of adaptation to occur in a fine-grained landscape.

Second, maladaptation in this system could arise due to underlying genetic differences in females preferring *T. arvense* to the native plant and differences in the larvae that fed on the toxic plant. Using a transcriptomic

approach, I elucidated that the females preferring *T. arvense* and the native host plant had similar gene expression profiles. However, gene expression differences between the larvae reflected host-plant specific changes suggesting signs of impaired feeding on *T. arvense*. Thus, I was able to identify that selection would likely act on larval ability to feed on *T. arvense* rather than females' ability to differentiate the hostplants.

Finally, the rate of adaptation and hostplant range could be related to the underlying rate of evolution and size of gene families related to detoxification enzymes along with those underlying gustatory and olfactory genes. I used a comparative phylogenomic approach to quantify gene family evolution of genes involved in plant compound detoxification and sensory genes between specialist and generalist brassicaceous feeders to address the causes of hostplant range and specialization. I found that gustatory and olfactory genes had a higher rate of gene evolution compared to other genes and genes related to detoxification evolved faster in generalist feeders compared to specialist feeders thus indicating that generalist feeders are more likely to incorporate new host plants in their diet compared to specialist feeders.

PREFACE

Local adaptation occurs when organisms respond adaptively in the presence of strong spatial heterogenous selection. It is known to produce phenotypes that have higher fitness in their native environment compared to any other phenotypes in the same environment. Although one would expect local adaptation to be ubiquitous, many organisms are (mal)adapted to their native environments. Of the many factors that can lead to maladaptation, rapid environmental change can lead to drastic decline in population mean fitness. Rapid environmental change can cause locally adaptive traits/ responses to fail, especially if the rate of adaptation lags or if populations do not possess the underlying genetic variation to adapt to their new environment. My dissertation aims to understand the mechanisms underlying maladaptation and its consequences in specialized species interactions.

I begin my dissertation by providing an overview of the literature on maladaptation. The first chapter is a literature review on understand the causes for persistent maladaptation. Although research on maladaptation has received interest in the last decade, the factors that cause persistent maladaptation, the underlying mechanisms that prevent organisms from escaping maladaptation and the outcome of persistent maladaptation is poorly understood. In the review, I explain the causes for persistent maladaptation with a special focus on

maladaptation in species interactions. I follow this up by providing examples of persistent maladaptation in species interactions and the role of evolutionary forces in shaping persistent maladaptation. I end this chapter by providing an account on studying maladaptation in species interactions and the role of genomics and comparative phylogenetics in helping us understand and predict the outcome of persistent maladaptation in species interactions. The rest of the chapters in my dissertation deal with understanding the persistent maladaptation between a native herbivore and its interaction with a lethal invasive plant.

In the second chapter, I discuss the role of interplay between evolutionary forces in maintaining a persistent ecological trap. I use the system of a native specialist herbivore (*Pieris macdunnoughii*) that lays eggs on a lethal invasive plant (*Thlaspi arvense*) which causes 100% mortality of the offspring. I use a whole genome sequencing approach to understand the role of population structure, gene flow, genetic variation, and selection in shaping persistent maladaptation. I found signatures of putative selection in the presence of high gene flow and low genetic variation in areas where *T. arvense* was present. This work prompted the next question on understanding the underlying mechanisms that prevented *P. macdunnoughii* from escaping the evolutionary trap and to dissect mechanisms in ovipositing adults as well as larvae to understand if preference-performance mismatches were leading to the maintenance of this ecological trap.

In the third chapter, I use transcriptomic approaches to test preference-performance mismatches, a phenomenon in which the oviposition preference of

the females is negatively correlated larval performance in maintaining the persistent maladaptation. I found no differential transcriptional activity in the females that chose one plant over the other, whereas I found host-plant specific signatures in gene expression of the larvae. Larvae that fed on the toxic plant showed upregulated transcriptional activity in genes related to immune system response to bacteria and those underlying lipid catabolism and ageing, suggesting plant toxin related changes in larval gut microbiome and impaired feeding. Larvae that fed on the native plant had genes related to metabolism, development and immune proliferation upregulated suggesting herbivore success. This work suggests that the signatures of selection I identified in chapter two is likely acting on larval ability to feed on the plant rather than females' ability to differentiate between the two plants based on transcriptomic differences in our study. This work also highlights the role of detoxification enzymes in the larvae in adapting to novel resources. Following this, I then investigated the evolution of detoxification and sensory genes in Lepidopteran insects feeding on Brassicaceae to understand the underlying mechanisms of hostplant adaptation.

In the final chapter, I use a comparative phylogenomic approach to understand the differences in detoxification and sensory gene evolution between specialist and generalist herbivores that feed on Brassicaceae. I found that generalists had large gene family expansions and contractions in detoxification and sensory genes compared to specialist herbivores. Additionally, I found that generalist herbivores had higher gene evolution rate in genes underlying general

detoxification processes while specialists had higher gene evolution in genes involved in detoxifying specific toxins. Finally, I found that nitrile specifier protein (NSP), a key innovation in Lepidoptera that enabled feeding on Brassicaceae was under pervasive purifying selection with few sites showing lineage specific selection. This work identifies the differences through which generalist and specialist herbivores process toxic compounds as well as providing an explanation to the constraints of evolution in specialized herbivores in adapting to new hostplants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

UNDERSTANDING PERSISTENT MALADAPTATION[1]

[1] Ravikanthachari N, Boggs CL. To be submitted to Biological Reviews.

Introduction:

Generations of ecologists and evolutionary biologists have employed the framework of adaptation to understand the power of natural selection in structuring biodiversity, and the interactions that shape our natural world (Boag & Grant, 1981; Colautti & Lau, 2015; Darwin, 1859; Grant & Grant, 2002, 2014; Hereford, 2009; Pelletier et al., 2009; Schluter et al., 2010; Vermeij & Vermeij, 1974; Watt et al., 2003; Williams, 1966). Although adaptation is expected to produce organisms residing near the peaks of their fitness landscape, several studies have demonstrated that organisms are at least one standard deviation away from their fitness peaks. (Austen et al., 2017; Barton & Partridge, 2000; Haller & Hendry, 2014). Most studies use the same adaptation framework to address the causes of these deviations, viewing them as an apparent "lack of adaptation". The number of studies that use the framework of maladaptation to understand the effects of natural selection on fitness are miniscule (Brady, Bolnick, Angert, et al., 2019; Brady, Bolnick, Barrett, et al., 2019). Recent reviews have therefore highlighted the need for focused attention on maladaptation to understand the causes of deviations from fitness peaks (Brady, Bolnick, Angert, et al., 2019; Diamond & Martin, 2020). Including a maladaptation framework to understand the underlying forces causing suboptimal organismal fitness will provide a comprehensive and nuanced understanding of the ecological and evolutionary processes that shape the diversity of life on earth (Brady, Bolnick, Angert, et al., 2019; Brady, Bolnick, Barrett, et al., 2019; Derry et al., 2019).

What constitutes maladaptation? As stated above, dismissing maladaptation as the converse of adaptation or simply the lack of adaptation hinders our ability to identify and elucidate the complex interplay between selection on, and the magnitude and direction of responses by, organismal fitness. Conceptually, defining maladaptation has been a challenge and hence there has been no consensus on what constitutes maladaptation. This, compounded with the lack of attention to maladaptation, has led to chaotic interpretations and confusion in understanding maladaptation. For example, definitions that consider that all life is maladapted to some degree (as a result of organisms residing away from their fitness peaks) introduces the same bias as using an adaptation framework to study fitness (Brady, Bolnick, Angert, et al., 2019; Crespi, 2000). Additionally, considering that the presence of selection is an indication of maladaptation (as a corollary to the presence of selection as an indication of adaptation) leads to a similar pitfall (Barton & Partridge, 2000; Brady, Bolnick, Angert, et al., 2019).

The definition of maladaptation can also vary depending on the scale at which it is addressed as well as the scope of the study. Brady *et al*., (2019) while recognizing the importance of these distinctions in addressing specific contexts of maladaptation, highlighted the need for a unification of relative and absolute maladaptation perspectives to address the gaps in our understanding of maladaptation. Relative maladaptation is often used by evolutionary biologists to define populations that have lower mean fitness compared to the mean fitness of other populations. Mean fitness in evolutionary terms focuses on changes in

genotype frequencies or traits while ignoring population sizes (Brady, Bolnick, Barrett, et al., 2019; Fisher, 1930; Haldane, 1924, 1932; D. Reznick, 2016). In contrast, absolute maladaptation is often used by ecologists to define declines in population sizes or extinctions without emphasis on the underlying genotypic frequencies or traits (Brady, Bolnick, Barrett, et al., 2019; MacArthur, 1984; Senner et al., 2017; Swift et al., 2017; Urban & Skelly, 2006).

Although there has been recent interest in maladaptation, most studies are limited to examining maladaptation only in response to climate change or across short time scales. Long term studies or those that investigate maladaptation in species interactions are scarce (Nakajima, 2014; Singer & Parmesan, 2019a; Steward & Boggs, 2020). Studying maladaptation in species interactions and incorporating long term monitoring will broaden our knowledge of the complex nature of maladaptation and helps us distinguish maladaptation from adaptation or transient maladaptation. Our focus in this review is to provide an overview on the causes of maladaptation in species interactions, highlight the need for studying persistent maladaptation and for approaches that could help researchers study persistent maladaptation in species interactions. To understand what constitutes as maladaptation, we build on Brady et al (year) and use the operational definition that maladaptation is reduced fitness of a trait, resulting in declines in population size of organisms in their native habitat compared to foreign habitats.

What causes of maladaptation?

Several recent reviews have provided detailed accounts on the mechanisms that cause maladaptation (Brady, Bolnick, Angert, et al., 2019; Capblancq et al., 2020; Cotto et al., 2019; Crespi, 2000; Diamond & Martin, 2020; Gougherty et al., 2021; Hansen et al., 2006; Magnan et al., 2016; Robertson & Blumstein, 2019; Thompson et al., 2002). We briefly mention the common causes and mechanisms that can lead to maladaptation.

Evolutionary processes as a cause of maladaptation:

Evolutionary processes including mutation, gene flow and balancing selection that lead to adaptation can also lead to maladaptation (Alleaume-Benharira et al., 2006; Crespi, 2000; Lenormand, 2002). Mutations, which can increase genetic variance and thus potential for adaptation, are maladaptive when they are deleterious or when they increase genetic load (Crow, 1970; Haldane, 1937). Mutations are more often deleterious than advantageous and thus cause maladaptation via mutational load (Jiang et al., 2011; Loewe & Hill, 2010; North et al., 2011; Sprouffske et al., 2018). An example is the case of Mueller's ratchet when deleterious mutations accumulate over time leading to maladaptation. This is often seen when recombination is low (Muller, 1950; Sprouffske et al., 2018).

Gene flow, while aiding the potential for adaptation by increasing standing genetic variation (Bossart & Scriber, 1995; Morrissey & De Kerckhove, 2009; Savolainen et al., 2007), can also dampen adaptation through outbreeding depression (Bolnick et al., 2008; Farkas et al., 2016; R. B. King & Lawson, 1995;

Savolainen et al., 2007). In the stick insect *Timema cristinae*, increased gene flow between populations in heterogeneous habitats was associated with increased maladaptation as it decreased the frequency of cryptic color morphs in the native site (Farkas et al., 2016; Montejo-Kovacevich et al., 2020).

While balancing selection leads to the persistence of phenotypic/genetic polymorphs and can be adaptive, it can also decrease the population mean fitness by increasing the frequency of phenotypes/alleles which are suboptimal. In common buzzards (*Buteo buteo*), the light and dark plumage color has lower fitness compared to the heterozygote intermediate morphs. However, the light and dark plumages are maintained in the population due to assortative mating, thus leading to maladaptation (Krüger et al., 2001).

Environmental perturbation as a cause for maladaptation:

Environmental changes can cause maladaptation in populations that are locally adapted. For example, increased anthropogenic perturbations of the environment have caused local and global population declines and extinctions (Ceballos & Ehrlich, 2002; Román-Palacios & Wiens, 2020; Valiente-Banuet et al., 2015; Winfree et al., 2009). The mean fitness of a population can decrease with fluctuating environmental conditions. If populations possess sufficient genetic variance to adapt to fluctuating conditions, local adaptation might occur (Bossart & Scriber, 1995; Fitzpatrick & Reid, 2019; Savolainen et al., 2007). However, if there is a lag in the response to environmental fluctuation or if populations do not possess genetic variation, maladaptation will occur (Brady,

Bolnick, Angert, et al., 2019). The strength of maladaptation is positively correlated with increasing unpredictability of the environmental variation as well as the increasing rates of environmental fluctuations (Nakajima et al., 2013; Robertson & Blumstein, 2019; Senner et al., 2017). In the migratory bird Hudsonian Godwit (*Limosa haemastica*), asynchrony due to climate change in certain populations delays the breeding of the birds by a week compared to the insect peak abundance, thus resulting in higher mortality of early chicks and declining bird population (Senner et al., 2017; Swift et al., 2017).

Rapidly fluctuating environmental conditions can result in a phenomenon of evolutionary traps, a type of maladaptation where organisms prefer low quality or lethal resources over higher quality resources due to mismatches in cue-preference (Brady, Bolnick, Angert, et al., 2019; Brady, Bolnick, Barrett, et al., 2019; Hendry & Gonzalez, 2008; Schlaepfer et al., 2002; Steward & Boggs, 2020). Populations can escape evolutionary traps if they have sufficient standing genetic variation that can lead adaptive cue responses (Hermisson & Pennings, 2005; Lande, 1975). However, the rate of adaptive responses to changes in the environment are often lagging. Thus, many populations are expected to be persistently maladapted leading to persistent evolutionary traps. Examples of evolutionary traps include aquatic insects that use polarization of the surface of water as ovipositional cues laying eggs on bright colored inanimate surfaces that produce increased polarized light (Black & Robertson, 2020; Lao et al., 2020), animals mistaking artificial sources of light as stars/moon for navigation cues (Hopkins et al., 2018; Robertson & Horváth, 2019; Szaz et al., 2015;

Witherington & Bjorndal, 1991) and insects ovipositing on exotic plants mistaking them for higher quality host plants (Nakajima et al., 2013; Steward et al., 2019b; Steward & Boggs, 2020). Environmental change in the form of degraded habitats can also lead to maladaptation. Organisms experiencing stabilizing selection in degraded habitats have a narrow fitness landscape, thus increasing the frequency of maladapted individuals in the population. Classic examples include declining populations of birds and mammals due to deforestation (Pimm & Askins, 1995; Semper-Pascual et al., 2019), declining species abundance in habitat sinks (Pulliam, 1988) and in range edges (Jump & Woodward, 2003; Vilà-Cabrera et al., 2019; Willi et al., 2018).

Maladaptation in species interactions:

The preceding examples illustrate the causes of maladaptation of species in response to changes in their abiotic environment or maladaptation due to changes in traits of the species. The following examples deal with the causes of maladaptation in species interactions. Maladaptation in species interactions is complex, as maladaptation in one of the partners in the interaction can have significant effects on the life history, ecology and evolution of the other species in the interaction (Castledine et al., 2020; Farkas et al., 2015; Thompson et al., 2002). Maladaptation in species interactions can arise due to the above-mentioned causes for single species, although the processes involved are often complex and dependent on the nature of the interactions (Nuismer et al., 2000; Thompson et al., 2002).

One of the major causes of maladaptation in species interactions is time lags created between the species involved in the interaction in the local community. Time lags can arise due to differences in relative rates of gene flow in the species involved (Adiba et al., 2010; Kaltz et al., 1999). In a host-parasite system, higher gene flow in the hosts causes maladaptation in the parasites and vice versa (Gandon et al., 1996). This is often due to a lag in adaptation of the species with reduced gene flow (Gandon et al., 1996; Gandon & Nuismer, 2009). In the interaction between the Canarian lizard (*Gallotia galloti*) and its blood parasite, the parasite has higher fitness and infection rates on non-native lizard hosts compared to native hosts. The lizards have higher rates of gene flow compared to the blood parasite, thus increasing genetic variation of the lizards in the native population which creates a lag in the adaptation of the blood parasite (Oppliger et al., 1999).

Maladaptation can also arise from varying selective pressures in antagonistic interactions. In host-parasite systems, the common host phenotype is often maladapted to the parasite compared to the rarer host phenotype as a result of frequency dependent selection (Forde et al., 2004; Siemens & Roy, 2005; Svensson & Connallon, 2019; Thompson et al., 2002). Density dependent selection can also influence the intensity of antagonistic interactions. For example, rapidly fluctuating hostplant densities can cause maladaptation in herbivores (Gering et al., 2019). In the butterfly *Pieris macdunnoughii*, increasing abundance of the exotic plant *Thlaspi arvense* can increase the magnitude of maladaptation as females have higher probability of laying eggs on the plant that

do not support larval growth (Nakajima et al., 2013; Steward et al., 2019b; Steward & Boggs, 2020).

Along with demographic factors and patterns, maladaptation also depends on the underlying genetic architecture of the species involved in the species interaction. In antagonistic interactions, higher genetic variance, ploidy, or high rates of recombination in one species can create time lags that cause maladaptation in the interacting partner. For example, plant species that are polyploid are often resistant to infestation and hence cause maladaptation in the parasites. Parasites whose underlying virulence is controlled by genes of major effect compared to genes of additive effects are better at adapting to hosts (K. C. King et al., 2012; Nuismer & Thompson, 2001; Oswald & Nuismer, 2007).

How does maladaptation vary in space and time in species interactions?

Heterogeneity in the landscape and temporal variation can influence the dynamics and trajectory of maladaptation in species interactions. Spatial heterogeneity can structure populations wherein subpopulations can experience varying levels of selection (Forde et al., 2004; Nuismer et al., 2000; Thompson, 1999a, 1999b). Each subpopulation can thus be adapted to the local conditions experienced. According to the geographic mosaic theory of coevolution, species interactions in space result in interaction-mediated mosaics of local maladaptation and adaptation (Nuismer et al., 2000; Thompson, 1999a).

Heterogeneity in the habitat can also influence patterns of gene flow and the strength of selection between/within metapopulations (Forde et al., 2007;

Garant et al., 2007). In a host-parasite system, maladaptation due to gene flow can occur in either the parasite or the host depending on relative local selection. For example, if the parasite is under selection in response to its host, gene flow in parasites can dampen local adaptation if the strength of the selection on the parasite is weak and local adaptation of hosts to their parasites can dampen if gene flow among host populations is high (Hoeksema & Forde, 2008). Gene flow between spatially heterogeneous habitats can also create lags between the interacting partners thus causing maladaptation in either some of the species in the interaction (in antagonistic interactions) or all the partners involved (in mutualisms) (Alleaume-Benharira et al., 2006; Forde et al., 2007; Gandon & Nuismer, 2009; Urban, 2006, 2011).

The strength of maladaptation due to spatial heterogeneity also depends on the type of species interactions. Specialized interactions are more sensitive to the dampening effects of gene flow from spatially varying habitats while generalist interactions are more buffered from the effects of gene flow. However, a study measuring the effects of spatial heterogeneity in a generalist mutualist system found that adaptation was influenced by spatial variation and previous selection regimes and populations that had experienced strong directional selection had higher fitness while populations that experienced less selection were generally maladapted. In specialized interactions, spatial heterogeneity can create varying degrees of maladaptation among the interacting partners. In the interactions between *Greya politella* and its host plants, in some populations the interactions were entirely antagonistic while in other populations they were

entirely mutualistic, and depended on the presence other pollinators in the habitats (Thompson, 1999b). Theoretical modelling of the interaction has shown that maladaptation was highest at the interface between mutualistic and antagonistic habitats, whereas adaptation was highest in habitats where mutualism was favored (Thompson et al., 2002).

Spatial heterogeneity can also create source-sink dynamics, thus leading to maladaptation and population declines (Delibes et al., 2001; Holt, 1996; Pulliam, 1988). The outcome of source-sink dynamics depends on the relative attractiveness of the habitat. If organisms have high preference for source habitats, population declines occur only if a threshold limit is crossed in the sink habitat (Delibes et al., 2001). If there is no preference for either the source or the sink, population decline is linear. If organisms prefer sink habitats compared to the source, then the population decline is steep (Delibes et al., 2001; Pulliam, 1988).

Temporal fluctuations can also cause maladaptation in species interactions. Maladaptation due to temporal variation can be transient or lasting. Unpredictable temporal fluctuations can cause time lags in species interactions leading to either maladaptation in all the species involved in the interactions (in mutualist systems), or a proportion of the species involved (in antagonistic systems) (Senner et al., 2017; Singer & Parmesan, 2010, 2019a; Swift et al., 2017). In mutualist systems, early snowmelt due to climate change causes asynchrony where plant growth and flowering precedes the emergence of pollinators, thus causing mismatches and maladaptation (Kudo & Cooper, 2019;

Kudo & Ida, 2013). In Hudsonian godwits, asynchrony in the breeding of the migratory birds and peak insect abundance causes higher mortality in the chicks (Senner et al., 2017; Swift et al., 2017). Temporal maladaptation can also occur at different time points during development. Organisms can be maladapted only during a particular life-stage (Cotto et al., 2019). For example, newly hatched turtles often mistake artificial lights as the light from moon and move towards the land instead moving towards the ocean thus causing higher mortality as a result of evolutionary traps (Witherington & Bjorndal, 1991). In some cases, even though organisms appear to be maladapted, it may represent natural cyclic variations of population contractions and expansions (Carlson & Quinn, 2007; Siepielski et al., 2017; Wissinger & Whiteman, 1992) and to infer maladaptation, long term studies are necessary.

How can organisms escape from maladaptation?

As described in the earlier sections, forces that can lead to maladaptation can also lead to adaptation and are hence, context dependent. Organisms can escape maladaptation through several routes. While gene flow in temporally stable habitats or from habitats with different selection mosaics can lead to maladaptation in the recipient population (Blanquart et al., 2013a), theoretical models have indicated that gene flow in temporally fluctuating environments increases standing genetic variation and thus adaptation and alleviates maladaptation (Blanquart et al., 2013a). Gene flow from populations that are pre-adapted to the recipient population's local conditions decreases maladaptation in the recipient population (Aitken & Whitlock, 2013; Caprio & Tabashnik, 1992;

Kelly & Phillips, 2016). For example, the populations of Tasmanian devils (*Sarcophilus harrisii*) are declining rapidly due to the contagious cancer, the Tasmanian devil facial tumor disease (DFTD). However certain populations in the northern part of the range of the Tasmanian devils are disease free as they are genetically resistant. Gene flow from these populations into regions where DFTD is prevalent can lead to adaptation and prevent population declines (Hollings et al., 2014; M. E. Jones et al., 2007). Maladaptation can also be alleviated if there is strong directional selection in response to local conditions (D. N. Reznick & Ghalambor, 2001; Shimada et al., 2011) (Nielson 2009, Shimada 2010, Reznick 2001). In a study investigating the prevalence of genes under directional selection in the Three-spined Stickleback (*Gasterosteus aculeatus*), physiologically important genes that have significant functions in adaptation to environmental heterogeneity were found to have a high incidence of directional selection compared to other genes (Shimada et al., 2011). In fact, when organisms experience strong directional selection, there is an initial population decline which may resemble maladaptation and is followed by population expansion when populations adapt to a new adaptive optimum (Bell, 2017; Bell & Gonzalez, 2011; Carlson et al., 2014). In small populations, if adaptive alleles arise in the population, genetic drift can spread the alleles through the population conferring adaptation. Apart from genetic drift, theoretical models have suggested that adaptive alleles spread faster in rapidly changing environments if they are expressed in heterozygotes (heterozygote advantage) (Haldane, 1924; Sellis et al., 2011). Maladaptation can also be alleviated through "rescues".

Populations can be rescued without adaptative genetic changes through demographic rescue (Hufbauer et al., 2015; Kanarek et al., 2014; Martin et al., 2000). In a simulation study using an eco-evolutionary model to understand the effect of immigration of pre-adapted populations in response to fishing vs immigration from unexploited population sources (sources with natural genetic variation unexposed to fishing) in preventing population declines in an Atlantic cod (*Gadus morhua*), immigration from unexploited population sources enhanced faster recovery compared to immigration from pre-adapted sources (Kuparinen & Uusi-Heikkilä, 2020). Individuals from surrounding sources immigrating into sinks can prevent absolute maladaptation (population extinction). Demographic rescue can also increase genetic variation or introduce new alleles to sweep through the population that are adaptive.

Another type of rescue, known as evolutionary rescue (Darwinian rescue) is a process that involves rapid adaptation through natural selection after an initial period of population decline due to maladaptation (Bell, 2017; Bell & Gonzalez, 2011). Using a laboratory yeast model and its response to salt stress to test the mechanisms underlying rapid evolution in a deteriorating environment, the frequency of evolutionary rescue was maximum when metapopulations of yeast experienced an initial gradual deterioration along with local dispersal (Bell & Gonzalez, 2011).

What causes persistent maladaptation?

Although much is known about the causes of maladaptation and routes by which organisms can overcome maladaptation, very little is known about what causes persistent maladaptation (Brady, Bolnick, Angert, et al., 2019; Brady, Bolnick, Barrett, et al., 2019; Derry et al., 2019). Several processes can lead to persistent maladaptation in species interactions. Broadly, persistent maladaptation can occur when adaptive evolutionary responses are insufficient to compensate for the negative effects of maladaptation (Bradshaw & McNeilly, 1991; Radchuk et al., 2019; Sinervo et al., 2010). This can arise when there is a lag in the rate of adaptation via natural selection in the underlying trait (Sinervo et al., 2010). Using historical resampling of 48 Mexican lizard from 200 sites, Sinervo et al. found that 12% of the populations had gone extinct and incidence of extinction was positively correlated with the rate of environmental change. Modeling of physiological and thermal preference responses were insufficient to adapt to changing temperatures and the threshold for response to climate change had been crossed and thus will lead to further extinctions of lizard species (Sinervo et al., 2010). Environmental perturbations, both abiotic as well as biotic may be faster compared to the rate of adaptation and thus lead to persistent maladaptation. Evolutionary responses might also be insufficient when there is low genetic diversity/variation in the focal population (Blanquart et al., 2013a). Without genetic variation for directional selection to act on, responses to selection may be insufficient. Trade-offs between traits or among correlation in traits can also result in insufficient responses to selection (Diamond & Martin, 2020). For example, adaptation to heat stress in ectotherms might have trade-

offs with fecundity or body size, thus constraining evolutionary responses to selection (Winchell et al., 2016).

Persistent maladaptation can also result from evolutionary processes. The interplay between gene flow and selection can either lead to local adaptation or transient or persistent maladaptation (Tufto, 2001). When directional selection is stronger compared to gene flow, local adaptation occurs. If gene flow is stronger, maladaptation might occur. However, if gene flow balances selection or if gene flow and selection are both weak, populations can experience persistent maladaptation. Mildly deleterious or weakly beneficial alleles can also result in persistent maladaptation by increasing the frequency suboptimal traits and thus decreasing the mean fitness of the population (Brady, Bolnick, Angert, et al., 2019). Since selection against mildly deleterious alleles or selection for beneficial alleles is weak, it can result in persistence of these traits and thus maladaptation. Frequency dependent selection can also lead to persistent maladaptation by increasing trait variance thus reducing fitness (Krüger et al., 2007). In antagonistic species interactions, trait mismatches arising from evolutionary constraints can lead to persistent maladaptation. In many species of Lepidoptera, female oviposition preference for a host plant and larval performance on the plant are often not correlated and females often make mistakes while laying eggs (Nylin & Janz, 1993). This can lead to persistent evolutionary traps when exotic species are introduced into the landscape. Females might have retained ancestral preference for plants that are similar to native plants and recognize the exotic host plants as potential host plants (Singer, 2021; Singer & Parmesan,

2010, 2019a; Steward et al., 2019b; Steward & Boggs, 2020). Unless selection to either adapt to the new plant or avoid the new plant evolves, maladaptation might continue to persist.

Maladaptation that might seem persistent might actually represent a trajectory of the population leading to extinction. When environmental change is rapid, pervasive and long lasting (Gomulkiewicz & Holt, 1995), maladaptation might represent a scenario of delayed extinction (Kuussaari et al., 2009). Many populations of ectotherms are experiencing maladaptation and steep population declines while some have gone extinct due to rising global temperatures. Maladaptation in such scenarios might represent inevitable extinction (Sinervo et al., 2010).

How does one study maladaptation?

Although studies that address maladaptation are minuscule compared to those that address adaptation, the insight gained from studying adaptation can be applied to studying maladaptation. Studying maladaptation is a two-step process.

The first step in the process is to identify and differentiate maladaptation from adaptation. The classical method to measure local adaptation is through reciprocal transplant experiments (Savolainen et al., 2007, 2013). In reciprocal transplant experiments, the focal population is locally adapted if the mean fitness of the focal population is higher in its native habitat compared to other population in the habitat. In a meta-analysis of studies of reciprocal transplant experiments ,

local adaptation was only present 70% of the time and undetected 30% of the time (Hereford, 2009). Undetected adaptation could represent maladaptation, which were ignored in those studies. Thus, employing multiple hypothesis testing in experimental design can be a straightforward way to identify maladaptation even in studies that specifically test adaptation.

Another important distinction is to define what constitutes adaptation and maladaptation. The null hypothesis when testing adaptation or maladaptation is that the performance of a phenotype on fitness is neutral. Adaptation occurs when the performance of a phenotype increases fitness and maladaptation occurs when the performance of a phenotype reduces fitness (Blanquart et al., 2013a). From an evolutionary perspective, the presence of balancing selection can be considered as evidence for maladaptation since natural selection combined with genetic drift leads to selective sweeps that in turn removes polymorphisms in the population (Charlesworth, 2006). This is also reflected at the genome level where regions that are under balancing selection or regions with high nucleotide diversity or excess polymorphisms are signatures of maladaptation as positive selection (and resultant adaptation) results in decreased nucleotide diversity and low polymorphisms.

The next step in studying maladaptation involves identifying the underlying causes of maladaptation. These are covered in the previous sections under causes for maladaptation. Along with identifying the proximal causes for maladaptation, distinctions between transient and persistent maladaptation while incorporating spatial heterogeneity are crucial to accurately understand the

magnitude and consequences of maladaptation. Using systems whose natural history is well known and systems that have been studied long-term are particularly useful to understand temporal and spatial variation of maladaptation. The methods used to elucidate the underlying mechanism and quantifying adaptation can be used to study maladaptation as well (Blanquart et al., 2013a; Capblancq et al., 2020) (Hoban et al., 2016; Stapley et al., 2010). The methods used to study maladaptation can be classified into theoretical models, ecological experiments, and genomic approaches.

Several theoretical modeling approaches such as genetic models including genetic additive models, models incorporating rate of gene flow and individual based models have been developed to study adaptation. Genetic models are most useful to understand relative maladaptation. Genetic models incorporating additive effects and effects of major genes are crucial in understanding the role of genetic drift and selection in adaptation (Lenormand, 2002). For example, models incorporating additive genetic effects of parasite adaptation to new hosts help us understand underlying general mechanisms that are likely to lead to adaptation on new hosts. Using this, one can also model the mechanisms that are likely to lead to maladaptation. For example, selection on? genes of major effect are likely to lead to adaptation while genes of additive effect often have suboptimal fitness (Lenormand, 2002; Yeaman & Otto, 2011; Yeaman & Whitlock, 2011). Understanding the models that incorporate constraints on gene evolution and genomic architecture can then provide clues as to the mechanisms that lead to maladaptation.

Models of gene flow have been used to understand adaptation as well as maladaptation. Gandon et al. (1996) developed a metapopulation model that examined the role of gene-for-gene co-evolution in local adaptation. They found that parasites were more likely to be adapted to their local hosts if they had higher migration rates compared to the hosts and postulated that the inverse should occur when hosts had higher migration rates. This model provides a basis for understanding how gene flow leads to maladaptation. The species with lower gene flow should be maladapted as reduced genetic variation might prevent it either from adapted to the local host (for parasites) or resist local parasites (for hosts).

Individual based models are crucial to understand maladaptation. Individual based models can provide clues on the mechanisms underlying absolute maladaptation. The butterfly *Pieris macdunnoughii* is maladapted to the exotic host plant *Thlaspi arvense* and the plant resembles an evolutionary trap. Individual based modeling describing movement of the butterfly inside the landscape of the evolutionary trap based on field data found that increasing patch shape, distance to the nearest native and the exotic plant and plant distribution caused declines in the population of the butterfly (Nakajima et al., 2013). Theoretical models which incorporate environmental heterogeneity have found that adaptation occurs when the change in the environment is along a steady gradient and organisms are unable to adapt when the environmental change is steep or near range edges (Bridle et al., 2019).

Ecological methods have been extensively used to study adaptation. Classical approaches including reciprocal transplant experiments, comparative analysis of trait fitness under different ecological conditions and phylogenetic comparative methods have been used extensively to understand adaptation. Several other reviews provide extensive information on using ecological experimental methods. Employing methods to understand absolute maladaptation can be accomplished by designing experiments to include multiple hypothesis testing.

Genomic methods including F_{st} based approaches, genotype-phenotype-environment approaches, selective sweep approaches, genetic mapping approaches and rates of non-synonymous substitutions in codons are used extensively to identify regions of the genome that are under selection (Garud et al., 2021; Gautier, 2015; Günther & Coop, 2013; Thornton et al., 2007). These methods are based on identifying underlying differences in allele frequencies between populations or contrasting phenotypes. These differences produce characteristic patterns of allele frequencies that can provide insight on adaptive and maladaptive processes. Adaptive processes that lead to fixation show presence of selective sweeps that include reduced nucleotide diversity, low rate of recombination, low levels of polymorphisms in the functional regions of traits that confer adaptation. Alternatively, functional regions of the genome that show signatures of balancing selection with increased polymorphism should be investigated as maladapted since adaptive selection is known to purge polymorphisms. GWAS approaches are specifically powerful in identifying

maladaptation as they can correlate underlying allele polymorphisms with phenotypes. Another recent approach known as natural co-GWAS can be specifically useful in understanding maladaptation in antagonistic species interactions (MacPherson et al., 2018). Natural Co-GWAS is a relatively new method that determines the statistical differences between each pair of host and parasite alleles and provides a graphical representation of the significant SNP associations in both hosts and parasites. Significant SNP correlations can provide evidence for regions that under co-evolution in both hosts and the parasites. However, reciprocal SNP correlations (selective sweep in one while balancing selection in the other) can provide clues on the interacting partner that is maladapted and hence can be a powerful tool to study maladaptation in species interactions. However, this method has only been tested in limited study systems (Märkle et al., 2021). Further development of natural co-GWAS can help us understand maladaptation in mutualistic interactions as well. co-GWAS coupled with long-term data can also shed light on understanding the trajectory of arms-race or trench warfare in co-evolution as well as understanding temporal variations in maladaptation.

How to study persistent maladaptation?

Long term demographic monitoring of populations can identify consistent declines that can warrant further investigation for persistent maladaptation. Long term monitoring is crucial in separating actual persistent maladaptation from transient or cyclic maladaptation. This provides critical information on factors and mechanisms underlying maladaptation which can be further tested using

transplant experiments or the genetic basis underlying maladaptation. A long-term study that extensively documented the interactions between *Euphydryas editha* and its introduced host plant *Plantago lanceolata* showed that organisms can become extinct when they adapt entirely to human-introduced resources and thus lose the potential to feed on native plants (Singer & Parmesan, 2019a). Since human introduced resources are unreliable both spatially and temporally, the populations are at a higher risk of maladaptation and extinction when the landscape changes with human activities (Nakajima et al., 2013; Singer & Parmesan, 2019a). Long term monitoring thus provides the first step in understanding if maladaptation is persistent.

Along with long term monitoring, theoretical models coupled with empirical studies and genomic studies help us understand the underlying mechanisms of persistent maladaptation.

Specific regions in the genome are hotspots for adaptation/selection while some are cold spots (Gomulkiewicz et al., 2007; Nuismer et al., 2000, 2003; Thompson, 1999a). Methods that incorporate spatial and temporal variation along with a population genomic approach (Vermeer et al., 2011) can identify and quantify these regions. Genomic sampling at periodic intervals can shed light into the extent of persistent maladaptation. Changes in allele frequency spectra can reflect changes in demographic history and changes in regions under selection. Persistent signatures of increased allele polymorphisms and increased nucleotide diversity in regions of the genomes underlying the corresponding trait can provide evidence for persistent maladaptation. Ultimately, understanding the

underlying gene-phenotype-environment effects of maladaptive genomic regions can facilitate understanding the causes for persistent maladaptation and to predict future maladaptation for the traits under investigation. Genetic methods that identify and quantify adaptation based on evolutionary processes including gene flow and selection can be used to understand persistent maladaptation. For example, Fst based methods such as BayPass (Gautier, 2015), BayEnv (Günther & Coop, 2013), EEMS (Al-Asadi et al., 2018), migrate-n (Beerli & Felsenstein, 2001) that can estimate gene flow and selection to test migration-selection balance can be used to understand how persistent maladaptation occurs in heterogeneous landscapes with varying selection mosaics. Furthermore, comparative genomic analysis between populations/individuals that vary in the degree of maladaptation and occupy different selection mosaics coupled with environmental data can help us identify environmental conditions and the underlying genetic mechanisms that are more likely to cause maladaptation. For example, as a result of climate change, many populations of species at lower elevations are maladapted to varying degree. Studying the mechanistic and genetic causes underlying maladaptation in lower elevation populations can help us predict the propensity of maladaptation in mid-elevation or higher elevation populations in the event of ongoing climate change. Genetic approaches that can incorporate evolutionary history along with allowing methods to predict maladaptation using sister species can help us understand if persistent maladaptation is predictable in species interactions especially in instances of evolutionary constraints, convergent evolution, or frequency

dependent selection. For example, in specialized herbivore-plant interactions, studying maladaptive hostplant use that occurs in one species can help us predict the risk of maladaptation in the sister species as host plant use is often conserved within clades. If maladaptive hostplant use has genetic underlying causes in the focal species, studying the spatial and temporal variation of the maladaptive hostplant in combination with evolution of the genes/genetic regions involved can help us identify molecular and functional differences/similarities in the sister species and in turn can be used to predict the risk of maladaptation in the sister species in the event of similar environmental conditions in the sister species.

Conclusion:

Given that maladaptation is ubiquitous, it has not been given the attention that adaptation has received. However, the recent influx of studies addressing the importance of maladaptation and its role in shaping the biotic world have increased the focus on studying maladaptation (Brady, Bolnick, Angert, et al., 2019; Brady, Bolnick, Barrett, et al., 2019; Cotto et al., 2019; Derry et al., 2019; Fitzpatrick & Reid, 2019). Recent studies, while acknowledging the importance of maladaptation, are concentrated on climate change-induced maladaptation (Capblancq et al., 2020). It is hence necessary to study maladaptation as a force by itself and not just through the lens of climate change. Studies which address several facets of maladaptation including maladaptation in species interactions and persistent maladaptation are crucial, as these help us understand ecological and evolutionary processes structuring communities and species' demographics

in nature. It is especially important to incorporate how maladaptation varies in heterogeneous habitats and temporally to better understand the effects and consequences of human-induced habitat change. A shift in experimental approach to include multiple hypothesis testing while designing adaptation studies can uncover underlying maladaptation in nature.

Several theoretical models as well as empirical approaches exist that are useful to understand maladaptation. These models can be further developed using long term ecological studies of maladaptation to build a framework to understand persistent maladaptation. Several genomic methods that are useful in understanding adaptation can also provide insights on maladaptation when combined with phenotypic fitness. Newer methods such as co-GWAS can revolutionize identifying and understanding the effects of maladaptation in species interactions when multiple hypothesis framework is included in such experiments. Comparative genomics and phylogenomic methods are extremely useful in helping us understand the role of habitat heterogeneity and temporal variation in shaping maladaptation in species interactions as well as persistent maladaptation. Finally, studying persistent maladaptation and/or maladaptation in species interactions is critical for identifying populations/species that are at a risk of population declines or extinctions. Understanding maladaptation in species interactions and maladaptation that is persistent can help us predict populations/species at risk due to anthropogenic introductions of non-native species and/or environmental change. Studying conditions that lead to persistent maladaptation can especially serve useful in evolutionary rescues in areas of

conservation and agriculture to prevent extinctions of endangered and at-risk species as well economic loss to due to extinctions of pollinators or crop plants.

Acknowledgements

We thank Ward Watt, Peter Andolfatto, Joseph M Quattro, Brian Hollis, Carrie Wessinger, Maria Stager, Jake Swanson, and Blair Stokes for providing comments on the earlier versions of the MS.

CHAPTER 2

LOCAL ADAPTATION OF A NATIVE HERBIVORE TO A LETHAL INVASIVE PLANT[2]

[2] Ravikanthachari, N, Steward RA, Boggs CL. To be submitted to Molecular Ecology.

Introduction:

Strong spatially heterogenous selection can lead to locally adapted populations. Local adaptation is maintained though an interplay of evolutionary forces, viz strong selection, and low gene flow between divergently adapted populations (Blanquart et al., 2013b; Hereford, 2009; Savolainen et al., 2013; Whitlock, 2015). The interplay between gene flow and selection can lead to a geographic mosaic with varying levels of adapted and maladapted sub-populations (Gomulkiewicz et al., 2000; Thompson, 1999a, 2005). The ability of the populations to adapt to local conditions depends on substantial genetic variation, larger effective population sizes and the predictability of seasonal and temporal changes in its native habitat (Franch-Gras et al., 2018; Hedrick, 2011; Lande, 1976; Lenormand, 2002; Perrier et al., 2020; Simons, 2011).

Novel environments can derail adaptative evolution by decreasing the fitness of individuals that are adapted to current local conditions (Brady, Bolnick, Angert, et al., 2019; Reed et al., 2003; Robertson & Hutto, 2006; Schlaepfer et al., 2002, 2005). Species experiencing novel environments in a part of their range can be locally maladapted, even though they may not experience any fitness loss in other parts of their ranges (Brady, Bolnick, Barrett, et al., 2019; Bridle et al., 2019). The outcome]interactions between native populations and novel environments depends strongly on population connectivity among divergently adapted populations as well as the adaptive potential to novel conditions in the population (Bolnick & Nosil, 2007; Farkas et al., 2015, 2016; Nosil et al., 2019). Identifying and understanding these interactions are critical to

understand species responses to novel environments, especially in the face of anthropogenic disturbance and climate change.

Herbivorous insects have a narrow range of hostplants that support development, and those insects form tightly co-evolved interactions with their hosts (Ehrlich & Raven, 1964; Hardy & Otto, 2014; Jaenike, 1990; Joshi & Thompson, 1995; Thompson, 1999b). Non-native or invasive species can breakdown long-standing interactions and lead to maladaptation in either or both of the interacting partners, either by outcompeting native species or through eco-evolutionary traps (LaForgia et al., 2020; Richard et al., 2019; Saul & Jeschke, 2015; Schlaepfer et al., 2005; Schweiger et al., 2010). Eco-evolutionary traps arise when previously adaptive traits become maladaptive in the face of novel resources or environments, such that the organisms prefer the low fitness-value novel resource instead of the high fitness-value native resources (Robertson et al., 2013; Robertson & Chalfoun, 2016; Robertson & Hutto, 2006; Schlaepfer et al., 2002).

Eco-evolutionary traps resulting from maladaptive oviposition choice have been documented in many insects (Horváth et al., 2007, 2010; Singer & Parmesan, 2018). Lepidoptera, specifically butterflies, are susceptible to laying eggs on low quality novel hosts plants (Nakajima & Boggs, 2015; Singer & Parmesan, 2010, 2019a; Steward et al., 2022; Steward & Boggs, 2020). Most butterflies are specialists in their hostplant choice and have co-evolved with their native hosts. These interactions break down when novel resources are introduced leading to maladaptation in the butterflies. Hence, they serve as an

ideal system to test the impacts of novel resources on local adaptation in species interactions. We here use a landscape genetics approach to identify genomic signatures of local adaptation in the specialist herbivore *Pieris macdunnoughii* Remington, 1954 (Pieridae; formerly *P. napi macdunnoughii*) (Chew & Watt, 2006) in the presence of the toxic invasive plant *Thlaspi arvense* (Brassicaceae). *Pieris macdunnoughii* females recognize *T. arvense* as a potential host plant due to chemical cue similarity with the native brassicaceous hosts and lay eggs on it (Chew, 1977b, 1977a, 1980; Nakajima, 2014; Steward et al., 2019b, 2022). However, larvae from eggs laid on *T. arvense* eventually die due to the toxicity of the plant. We sampled *P. macdunnoughii* in the East River Valley, Gunnison County, Colorado, where the butterfly oviposits on both native host plants and *T. arvense*.

Using genome-wide data from whole genome resequencing of individuals from areas where both the plants occur and from areas where *T. arvense* is absent, we identified candidate single nucleotide polymorphisms (SNPs) that are under selection in relation to the presence of *T. arvense* in the habitat. We specifically quantify: i) population structure of *P. macdunnoughii* in the East River Valley, ii) gene flow patterns of *P. macdunnoughii* between areas with and without *T. arvense* and iii) underlying signatures of selection in *P. macdunnoughii* in response to the presence of *P. macdunnoughii* in the landscape. Our work builds on decades of existing research on this system to address the genetic basis underlying maladaptive behavior in the face of novel resources.

Materials and methods:

Study organisms

Pieris macdunnoughii is a montane butterfly distributed across the southern Rocky Mountains in North America. *Pieris macdunnoughii* is a specialist herbivore, and females oviposit on and the larvae feed on several native Brassicaceae (mustard) plants including *Boechera spp*., *Cardamine cordifolia*, *Draba aurea* and *Descurainia incana* (Chew, 1977b, 1977a, 1980). Pierini larvae have evolved resistance to the toxicity of glucosinolates (secondary metabolites) found in the Brassicaceae by modifying the glucosinolate hydrolysis pathway to form thiocyanates and epithionitrilies that are less toxic, instead of forming isothiocyanates (Edger et al., 2015; Wheat et al., 2007). Although Pierini butterflies are resistant to glucosinolate toxicity, they are susceptible to species whose glucosinolate concentration or composition differs from the ones with which they have co-evolved locally.

Thlaspi arvense (L.) (Brassicaceae) is native to Eurasia and was introduced to Gunnison basin, Colorado in 1880s (Chew, 1977b). It is an early successional plant and is found in heavily disturbed areas. It has colonized elevations up to 2900m (Best & McIntyre, 1975; Warwick et al., 2011). *Pieris macdunnoughii* females recognize *T. arvense* as a potential host plant in areas where they both co-occur due to cue similarity and lays eggs on it. However, none of the larvae reared on *T. arvense* survive, thus leading to an evolutionary trap (Nakajima, 2014; Steward & Boggs, 2020). Previous research has indicated that the chemical profile of *T. arvense* is comprised mostly of the aliphatic glucosinolate sinigrin, compared to native mustards (e.g., *C. cordifolia*) that

contain both aliphatic and aromatic glucosinolates (Rodman & Chew, 1980). Additionally, recent research has indicated that sinigrin acts as a pre-ingestive deterrent and larvae exhibit delayed feeding on *T. arvense* compared to *C. cordifolia* (Steward et al., 2019b).

Population sampling and study area

The study area is situated in the East River Valley (ERV), Gunnison County, Colorado spanning an altitude from ca. 2800 to 3400m a.s.l. (fig. 2.1). We sampled 100 female adult butterflies during June - August 2016 from 5 locations and 235 adult butterflies of both sexes during June - August 2019 from 7 locations. Our final sample size consisted of 335 individuals from 11 locations. The distance between sampling sites ranged from 1km to 28km. All butterflies were caught using an aerial net. We collected whole body specimens of females and stored them in 98% ethanol in 2016. In 2019, 2mm x 2mm square of hindwings and a pair of midlegs were clipped and stored in 98% ethanol, after which the butterflies were marked and released. We also surveyed the presence/absence of *T. arvense* in all our sampling locations. We recorded *T. arvense* as present in 2 sites in 2016 (of 5 sites sampled) and 2019 (of 7 sites sampled). Our final sampling list consisted of 3 sites where *T. arvense* was present (of 11 sites sampled).

DNA extraction and WGS library preparation

We extracted genomic DNA from thorax (2016 samples) and from wing clips and a pair of midlegs (2019 samples) using the DNeasy Blood & Tissue Kit

(Qiagen, USA). We followed the manufacturer's protocol, with the following modifications: We increased the Proteinase K incubation step to 16 hours, used ice-cold ethanol to precipitate DNA, and heated the elution buffer to 57 ºC before DNA elution. We quantified the DNA concentration using a NanoDrop ™ 2000/2000c. We used an in-house Tn5 tagmentation protocol following (Andolfatto et al., 2011) for our library preparation. 2-10ng/ul of genomic DNA was tagmented with 1:4 diluted Tn5 transposase and 40uM pre-annealed oligonucleotides. The tagmentation was carried out in a final volume of 20uL containing 4uL of 5x TAPS buffer. Samples were incubated at 55 ºC for 7 minutes followed by rapidly lowering the holding temperature to 10 ºC. The Tn5 transposase was deactivated using 2.5uL 0.2% SDS at 55 ºC for 7 minutes with a heated lid. PCR-based barcoding and enrichment was carried out in a reaction mixture of 20uL consisting of 2 uL of the tagmentation reaction product, 10uL of 2x OneTaq Hot Start DNA polymerase, 4uL of nuclease free water and 2uL (10uM) each of i7 and i5 Illumina primer. The thermocycling conditions included initial denaturation at 72ºC for 3 minutes followed by denaturation at 94ºC, followed by 10 cycles with 10s at 94ºC, 15s at 62ºC, 30s at 68ºC and final extension for 5min at 68ºC. We constructed 5 replicate libraries for each sample, then pooled all the replicates and samples (1920 libraries). Size selection of 300-500 bp was carried out using 0.8x volume and 0.6x volume AmpureXP beads. All the libraries were sequenced at MedGenome on an Illumina Novaseq S4 platform, generating 150bp paired end reads.

SNP (Single nucleotide polymorphisms) calling

Demultiplexed raw Illumina reads and adapter sequences were trimmed using fastp (Chen et al., 2018) and mapped against the reference genome (Steward et al., 2021) using bwa mem with default parameters (Li, 2013). The resulting sam files were converted to bam format, sorted and indexed using samtools (Li et al., 2009b). Duplicates were marked and identifier groups were added using PICARD TOOLS with default parameters ("Picard Toolkit," 2019). Single nucleotide polymorphisms (SNPs) were called across all samples using GATK HaplotypeCaller to generate individual intermediate gVCF files that were then imported using GATK GenomicsDBImport and were finally genotyped using GATK GenotypeGVCFs. We used GATK VariantFilteration to further hard filter SNPs and Indels separately using the best recommended workflow practices (Van der Auwera et al., 2013), and finally used GATK SelectVariants to include only those variants that met the criteria: QD > 2, base quality > 30, SOR < 3.0, FS < 60, MQ > 40, MQ Rank Sum > 12.5, Read Pos Rank Sum > 8.0 . We then used VCFtools to retain only high quality bi-allelic variants using the following parameters: minimum allele count = 4, max missing = 0.20, min q= 30, min-mean DP = 6, max mean DP = 60, minDP = 6, max DP = 60. We further used PLINK (Purcell et al., 2007) to prune those loci that were at LD with the following parameters: r^2>0.2 in a window of 50bp. Our final sample size consisted of 335 individuals and 768,339 SNPs.

Population structure

We used a subset of 735,000 putatively neutral SNPs that were obtained after removing outlier loci (see below) to discern the population structure.

Principal Component Analysis (PCA) clustering was carried out using PLINK (Purcell et al., 2007). We complemented our PCA analysis with archetypal analysis following (Gimbernat-Mayol et al., 2022), to test biases due to irregular sample sizes with PCA analysis and to identify latent factors. We removed multi-allelic SNPs prior to running archetypal analysis and performed the analysis with k ranging from 2 to 4. We picked the one whose PC1 and PC2 axes best accounted for the variance. We additionally used ADMIXTURE to corroborate the results from PC and archetypal analysis and to assess population structure and ancestry (Alexander et al., 2009). We performed ADMIXTURE analysis for ancestral clusters K ranging from 1 to 6 and selected the K value with the lowest cross-entropy as the best estimate of population admixture.

Nucleotide summary statistics

We used geoVar (Biddanda et al., 2020) to estimate allele frequency distribution (site frequency spectrum) in a) all 11 sites in the East River Valley and b) in areas with and without *T. arvense*. We used the allele frequency distribution in geoVar in place of widely used SFS (Site Frequency Spectrum) methods (Gutenkunst et al., 2009), since geoVar permits simultaneous comparisons of SFS for more than 2 populations. Furthermore, geoVar allows us to classify minor allele frequencies into common (>5% frequency among all samples), low (1 - 5% frequency), rare (<1% frequency) and unobserved (allele not present). This approach helps us to decipher patterns of genetic variation within and among populations that are not possible by other methods.

We first removed all multi-allelic sites in our VCF and converted it into a frequency table using geoVar and calculated the cumulative fraction of variants that contributed to the allele frequency pool for each site/habitat. Using VCFtools (Danecek et al., 2011), we calculated expected and observed heterozygosity of butterflies for sites where *T. arvense* is present and sites where *T. arvense* is absent. We used pixy (Korunes & Samuk, 2021) to estimate genome-wide nucleotide (π) diversity. For the input for pixy, we used both the invariant and variant sites in our input VCF file as recommended by the authors to prevent bias in our estimates. We used a 10kb sliding window with a 50bp step to calculate π estimates. We used VCFtools to calculate pairwise Fst for all combinations of sites in the East River Valley. We used a sliding window approach with an interval window of 1kb with a 50bp step for estimating pairwise Fst.

Migration surface/gene flow

We used divMigrate from the diveRsity package (Keenan et al., 2013) in R (R Core Team, 2020) to estimate directional relative migration rates among sampling sites and PGDSpider (Lischer & Excoffier, 2012) to process the input for divMigrate. The relative migration network was scaled to the largest estimated magnitude based on N_m as a measure of genetic distance. We assessed the significance of the migration network by running 1000 bootstrap iterations.

We used EEMS (Petkova et al., 2015) to examine spatial variation in migration among populations (demes) and genetic diversity within populations. EEMS estimates genetic differentiation using an isolation-by-distance model of

geo-referenced samples to visualize patterns of potential barriers and corridors of gene flow. We used PLINK to convert the VCF file to BED files and calculated a genetic dissimilarity matrix using bed2diff in EEMS. We ran EEMS using 400 and 800 demes and a MCMC run with 1.5 million burn-in iterations followed by 15 million sampling iterations. We ran multiple iterations and adjusted the proposal variance rates after each run till our runs converged (Fig S2.1). The parameters used for the final run are provided in the supplementary file (Table S2.1). We then combined the final output of all the demes to produce a composite migration and diversity landscape. Migration and diversity rates were illustrated on a log10 scale relative to the overall migration and diversity across the entire landscape, such that a rate of one indicated tenfold higher migration and diversity rates relative to the average.

Genotype-environment association (GEA) analysis

Genome-wide scans for outlier loci were carried out in BayPass (Gautier, 2015). We first used the standard covariate model with the Importance Sampling (IS) approximation (-covis) in BayPass to obtain the following parameters: a) the population scaled covariance matrix and b) the XTX (SNP-specific Fst that corrects for observed population covariance) score of overall differentiation among sites. We then used the auxiliary covariate model using the MCMC approximation along with the Bayesian auxiliary variable to identify loci that were associated with the presence/absence of *T. arvense*. We additionally simulated 100,000 loci using the BayPass sim to calibrate the top 1% significance threshold for XTX. For environmental association, we used the Bayes Factor (BF) metric in

deciban units (dB) as a measure of support for association with *T. arvense*. We used a cut-off of db>20 (decisive evidence) in favor of association based on Jefferey's rule (Jeffreys, 1961). Loci associated with *T. arvense* as well as those under adaptive differentiation were used as input in SNPEff (Cingolani et al., 2012) and BlastX (Altschul et al., 1990) to identify the underlying genes as well as their effects on protein structure.

Results:

Population structure and admixture

Our Principal Component Analysis (PCA) and archetypal analysis using neutral loci failed to separate individuals into distinct clusters and hence did not reveal any population structure among all sites in the East River Valley (PCA: PC1=6.56% and PC2=5.65%; fig. 2.2a; archetypal analysis: PC1=52.92%, PC2=47.08%; fig. 2.2b). Our ADMIXTURE analysis corroborated the results of the PCA and archetypal analysis. The lowest Cross Validation (CV) indicated K =1 (supplementary table 1), thus indicating that sites with *P. macdunnoughii* across the ERV comprise a single population.

Genetic variation and differentiation

The allele frequency distribution for *P. macdunnoughii* in habitats with and without *T. arvense* were mostly low frequency alleles that were localized to each habitat (fig. 2.3a; 26% of cumulative fraction of variants), followed by 22% of shared and common alleles. The rest of the variants were either rare or low frequency alleles that were localized to each habitat (fig. 2.3a). In contrast to the

habitat-based allele frequency distribution, pairwise comparison of allele frequencies among sites consisted of alleles that were shared amongst all the sites (fig. 2.3b).

We found higher genetic variation among *P. macdunnoughii* in areas without *T. arvense* compared to areas with *T. arvense* (fig. 2.4a; Bartlett's K2=6.95; df=1; p=0.008). The observed heterozygosity was less than expected in areas without *T. arvense* (fig. 2.4a; Bartlett's K2=39.63; df=1; p<0.001) and the observed vs expected heterozygosity was similar in areas with *T. arvense*. Our genome-wide average estimates of π (nucleotide) diversity indicated a significant increase in nucleotide diversity in areas without *T. arvense* compared to areas with *T. arvense* (fig. 2.4b; Bartlett's K=9.611; df=1; p=0.001).

Our genome-wide mean pairwise Fst among all combination of sites was less then <0.001 (Fig S2.2), thus suggesting low genetic differentiation and high relatedness among sites. Our correlation map based on the population co-variance matrix indicated that all sites were weakly positively correlated. Finally, the Mantel test did not reveal any signs of isolation by distance (p=0.14, r=0.05).

Migration

We observed significant variation in the magnitude and direction of relative migration of *P. macdunnoughii* among all sites (fig. 2.5). The estimated relative migration rates ranged from 0 to 1 with an average of 0.56. We observed unidirectional and bidirectional migration from areas with *T. arvense* to areas without *T. arvense* and vice versa (fig. 2.5). The highest rates of gene flow

(Nm>0.85) were observed from Gothic Townsite to Elko Park (Nm=0.89) and vice versa (Nm=1), and from Quigley Creek to Gothic Townsite (Nm=0.96). Additionally, we observed high rates of relative gene flow to and from Gothic Townsite compared to other sites (fig. 2.5).

Estimated effective migration surface (EEMS) contours revealed the low relative effective gene flow of *P. macdunnoughii* in the East River Valley (fig. 2.6a). Most of the potential barriers also showed high posterior probabilities in the Bayesian estimation of migration parameters (Fig S2.3). Confidence intervals of migration parameters for sites in the Upper ERV (Copper Creek, Copper Creek 1st Crossing, Gothic Townsite, Rustler's Gulch, Quigley Creek, and Elko Park) were higher (>0.95) compared to the sites in the Lower ERV (Lower and Upper Brush Creek and Lower and Upper Cement Creek) had lower confidence <.90. We also estimated relative effective genetic diversity of *P. macdunnoughii* in the ERV. Our analysis highlighted low relative effective genetic diversity of *P. macdunnoughii* in all sampled sites (fig. 2.6b). All sites showed a posterior probability of >.90 (Fig S2.4).

Genotype-environment association (GEA) analysis

The BayPass core model (-covis) allowed us to estimate the scaled covariance matrix of population allele frequencies that quantify genetic relationship among pairs of sites. The results of Ω estimates were in agreement with our F_{st} estimates indicating that all sites are genetically similar (Fig S2.2).

For the identification of outlier loci, the XtX (SNP specific F_{st}) estimates were calibrated by analyzing a POD (pseudo-observed data set) of 100,000 SNPs. At the 1% threshold (XtX >26) for POD, we identified 8600 outlier SNPs (Fig S2.5). Our analysis that included presence/absence of *T. arvense* (as the environmental covariate) under the auxiliary covariate model identified 1008 SNPs that were associated with the presence of *T. arvense* (BF >20) (Fig S2.6). Overall, we identified nine SNPs that were shared by both the XtX outlier loci analysis and environmental association analysis (fig. 2.7). Our SNPeff analysis indicated that the nine SNPs caused 25 functional effects in protein coding genes. Of these 25 effects, one was classified as low impact (unlikely to change protein behavior) and 24 were classified as modifiers (effects on non-coding regions). The BlastX analysis of these regions revealed that genes underlying larval development and metabolism were under selection and associated with presence of *T. arvense* (Table 2.1).

Genic environment of outlier loci

The nine outlier SNPs were distributed across five chromosomes and one unplaced scaffold (Table 2.1), including one on the sex chromosome. The SNPs had 25 effects on their surrounding genic environment. The location of the 25 effects were as follows: downstream of a gene= 8, exonic region=1, intergenic region=4, intronic region=5, transcript=5 and upstream of a gene=2.

Discussion

Pieris macdunnoughii in the East River Valley comprise a single admixed population. Allele frequency differences between habitats classes (presence/absence of *Thlaspi arvense*) were largely driven by localized alleles, whereas pairwise allele frequency differences among sites were driven by geographically widespread alleles. Additionally, heterozygosity of butterflies was lower in areas where *T. arvense* occurred compared to areas where the plant was absent. Examining relative and effective gene flow patterns revealed bidirectional, asymmetric relative gene flow and low effective gene flow across the ERV. Finally, we identified signatures of selection in *P. macdunnoughii* in response to *T. arvense*. That selection likely occurred in larval ability to feed on T. arvense rather than the females' ability to differentiate between potential hosts. In sum, we identified ontogenic-specific genetic signatures of local adaptation in a native herbivore in response to a novel toxic hostplant at a fine geographic scale in the absence of strong population structure and genetic differentiation. Our work also expands the existing literature on herbivore responses to novel resources by combining landscape genomic and functional gene annotation approaches to dissect specific ontological stages where local adaptation was more likely to occur.

Evidence for a single population among sites.

Our results did not separate individuals into distinct clusters (fig. 2.1a, b) and indicated that all sites comprise a single population. Our results are in line with previous mark-recapture estimates of *P. macdunnoughii* in the 1970s and early 2000s (Nakajima, 2014). *P. macdunnoughii* occurs across elevations up to to

4500m.a.s.l and disperses along large elevational gradients (Nakajima, 2014). Several species of *Pieris* are known to undergo seasonal migration thus increasing gene flow and blurring population boundaries (Kitahara, 2016; Spieth & Cordes, 2012; Takami et al., 2004).

Distinct allele frequency distribution patterns between habitats and among sites.

Recent research has highlighted the utility of AFS/SFS to identify outlier loci, understand population structure, infer demographic changes as well as identify signatures of positive selection. In our study, AFS differences of *P. macdunnoughii* between areas with and without *T. arvense* was largely driven by alleles that were localized to each habitat even in the absence of strong population structure. However, overall differences among sites were largely due to differences in shared, common alleles. This is in contrast with theoretical and empirical work in other systems, which showed that differences between populations were driven mostly by localized low/rare frequency alleles and pairwise differences among sites were driven by shared, common alleles (Biddanda et al., 2020; Gutenkunst et al., 2009). For instance, analysis of large-scale human genomic data from the 1000 Genomes Project (1KGP) showed that allele differences between populations were due to localized and rare alleles and differences between individuals regardless of population origin were due to common variants found globally (Biddanda et al., 2020). Allele frequency differences between populations can result due to local adaptation and/or new mutations specific to the population (Günther & Coop, 2013). In our case, the

abundance of localized, low frequency alleles between the habitats could be due to the selection imposed by *T. arvense* leading to putative adaptive alleles.

Mechanism underlying local adaptation with high gene flow and low genetic variation.

Our results highlight important mechanisms that underly local adaptation. Population genetics theory suggests that local adaptation occurs in the absence of gene flow and that high gene flow often leads to maladaptation (Bachmann et al., 2020; Farkas et al., 2016; Gandon et al., 1996; Garant et al., 2007; Kirkpatrick & Barton, 1997; Lenormand, 2002). However, recent studies showed that local adaptation occurs in the presence of strong gene flow in certain scenarios (Fitzpatrick et al., 2015; Tigano & Friesen, 2016). For example, in spatially and temporally varying habitats, gene flow augments standing genetic variation and thus increases local adaptation (Blanquart et al., 2012, 2013b). Local adaptation can also occur in the presence of gene flow through adaptive introgression and strong selection against the immigrant alleles (Griffiths et al., 2021; Leroy et al., 2020; Rendón-Anaya et al., 2021). Our results indicated that all sites in the East River Valley were panmictic, with high gene flow, low genetic differentiation, and low genetic variation within the population. Nonetheless, we were able to identify strong signatures of local adaptation in *P. macdunnoughii* where *T. arvense* occurred. Eggs laid on *T. arvense* eventually die and thus there is strong selection against immigrant individuals that prefer *T. arvense* (Nakajima, 2014; Nakajima & Boggs, 2015). Similarly, local adaptation in lava flow lizards (melanism) occurred in the presence of strong gene flow and low genetic variation (Krohn et

al., 2019). This is in line with recent theoretical and empirical work that suggests that environmentally driven local adaptation does not lead to genome wide differences or require substantial standing genetic variation unless the underlying traits are linked to reproduction (Feder et al., 2012; Krohn et al., 2019; Shafer & Wolf, 2013).

Selection on oviposition vs larval performance.

Targets of selection in herbivorous insects' response to novel hostplants is likely polygenic, as it involves genes underlying oviposition, larval feeding, larval metabolism, and detoxification, as seen in several studies that have identified the genetic basis of oviposition or larval performance in insects (Egan et al., 2015; Gompert et al., 2022; Vertacnik & Linnen, 2017). We identified 9 loci distributed across five chromosomes, causing 25 variant effects. Identification using SNPeff and BlastX revealed that the majority of the genes were involved in larval development and metabolism and most of the variant changes affected non-coding regions. Non-coding regions in the genome contain regulatory elements that play a critical role in protein assembly, gene expression and regulation and are under purifying selection (Andolfatto, 2005; Bird et al., 2006; Loehlin et al., 2010). Therefore, these variant changes might have significant effects in larval feeding and detoxification. We did not identify any sensory (olfactory, gustatory, or visual) genes under selection that underlie female oviposition choice (de Fouchier et al., 2017; Engsontia et al., 2014; Ramaswamy et al., 1987; Yang et al., 2020). Lepidopteran females use a combination of sensory receptors to identify a potential hostplant (Haverkamp et al., 2018; Renwick & Chew, 1994;

Thompson & Pellmyr, 1991). Females use olfaction and visual cues for long range detection of hostplants and the final decision is made after gustatory tactile contact with the hostplant using the first pair of foretarsi in their legs (Ozaki et al., 2011; Ryuda et al., 2013). Differences in oviposition choice in females are driven by underlying differences in their chemosensory repertoire. Thus, our failure to identify chemosensory genes in our outlier analysis suggests that selection might be acting on larvae instead of the females. Existing evidence suggests that error prone oviposition in Lepidoptera females is the fuel driving hostplant range expansions and adaptation to new hosts, since repeated oviposition on less suitable/non hostplants imposes selection on the larvae to evolve to feed on the plant (Janz et al., 1994; Nylin et al., 2000; Nylin & Janz, 2009; Stefanescu et al., 2012).

Evolutionary constraints on oviposition preference and larval performance

Pieris macdunnoughii and its sister taxa in North America are derived from the Eurasian *P. napi* during the last Holarctic speciation event (Chew & Watt, 2006). *Pieris napi* larvae can develop successfully on *T. arvense* and other invasive Eurasian mustards in the Eastern US including *Alliaria petiolata* (Forsberg, 1987; Friberg & Wiklund, 2019; Prasai & Karlsson, 2011). *Pieris oleracea*, another species in the *Pieris* species complex is unable to use *A. petiolata* due to its toxicity to the larvae, in a case similar to that of *P. macdunnoughii* and *T. arvense* (Chew, 1977b; Haribal et al., 2001, 2001; Haribal & Renwick, 1998; Huang et al., 1994; Keeler & Chew, 2008). Thus, the North American *Peiris* larvae have lost the ability to develop on introduced mustards from Eurasia

while females have retained the ancestral preference for hostplants. This mismatch in preference-performance is the underlying cause for maladaptation in the larvae. Our results suggest that any adaptation by the butterfly to be able to use the plant will likely involve evolution of larval ability to develop on *T. arvense* and not of the females' avoidance of oviposition on *T. arvense*. Adaptation in response to *T. arvense* in the larvae would involve longer survival on *T. arvense*. Indeed, recent research has shown that after decades of maladaptation of *P. oleracea* on *A. petiolata*, the larvae are now able to develop on the plant during its bolting stage but not the rosette stage (Keeler & Chew, 2008). This highlights the importance of understanding the interactions between plant phenology, plant chemical composition and larval performance. Our results suggest that a similar outcome in *P. macdunnoughii* in the event of rapid evolution due to selection. Future research quantifying fine scale spatial and temporal patterns of larval performance coupled with quantifying *T. arvense* distribution and variation in plant chemical profiles might provide insights on the escape from or persistence of the evolutionary traps.

Conclusions:

Instances of interactions between native and non-native species are projected to increase across the globe due to range shifts, competition from invasive species and climate change. Understanding the effects of these interactions requires careful dissection of the ecological and molecular processes that mediate these interactions. Our results build upon decades of research aimed at understanding the causes and maintenance of evolutionary traps, to

elucidate the molecular responses and potential adaptation to novel resources. As insects are currently experiencing a global decline due to various factors including the ones mentioned above (Hallmann et al., 2017; Nakajima & Boggs, 2015; Wagner et al., 2021), the *Pieris macdunnoughii* - *Thlaspi arvense* system can serve as a model to understand and predict the outcomes of these interactions in insects even at a fine geographical scale. Our work adds to the growing literature aimed at understanding the impacts of novel resources on native species interactions and thus predicting the outcome of these interactions (Brady, Bolnick, Angert, et al., 2019; Farkas et al., 2015; Gilroy & Sutherland, 2007; Robertson et al., 2013; Schlaepfer et al., 2002). This work builds upon decades of work on the *Pieris*-*Thlaspi* maladaptation system and identifies the molecular mechanisms that underlies the evolutionary trap, thus pioneering it as a model system to understand maladaptation and evolutionary traps in the face of climate change (Chew, 1977b; Nakajima, 2014; Steward et al., 2019b, 2021, 2022; Steward & Boggs, 2020). Our work also highlights the potential of adaptation to occur in a fine-grained landscape in the absence of genetic variation and high gene flow.

Acknowledgements:

We thank W. Watt, J. Quattro, P. Andolfatto, B. Hollis and C. Wessinger for providing feedback on the manuscript. We thank J. Nicklas, L. Tran, L. Fisher, L. Hubbard, T. Feemster, M. Brown, M. Olson, S. McDaniel, H.Walton, E. Wagner, and O. Shrestha for assisting with field work. We thank K. Deitz and A. Pinharanda for assistance and advice on Tn5 library preparation. We thank K. Correa Veléz, J. Rivera-Hernandez, B. Stokes, J. Stewart, and D. Reisman for advice on PCR and sequencing. M. Kuwalekar, S. Baral,

R. Deshmukh, M. Stager, O. Lindestad, D. Speiser and C. Wessinger provided feedback on bioinformatic data analysis. Field research permits from the US Forest Service were arranged by the Rocky Mountain Biological Laboratory (RMBL). Funding was provided by the University of South Carolina College of Arts & Sciences to CLB, an NSF Graduate Research Fellowship to RAS, and RMBL graduate fellowships to NR and RAS.

Table 2.1: Annotation of outlier loci identified by BayPass XtX and BF (dB).

Chromosome	Position start	Position end	Codon	BlastX identification
1	153581	153583	STOP	cilia- and flagella-associated protein 410 isofrom
7	3428853	3428855	STOP	uncharacterized protein
9	11545167	11545169	START	ARFRP 1
9	4592359	4592361	START	homeobox protein Hox-A3-like
9	4597927	4597929	START	retinal homeobox protein Rax-like
9	4595024	4595026	STOP	homeobox protein Hox-A3-like
9	4599618	4599620	STOP	retinal homeobox protein Rax-like
14	11697968	11697970	START	DOCK 4
14	5730927	5730929	START	mediator of RNA polymerase II transcription subunit 15
14	11701979	11701981	STOP	inx3
14	5742219	5742221	STOP	tripartite motif-containing protein 45
xfSc0000009	573908	573910	START	uncharacterized protein
xfSc0000009	261794	261796	STOP	protein ALP1-like
xfSc0000009	263610	263612	STOP	Transposase

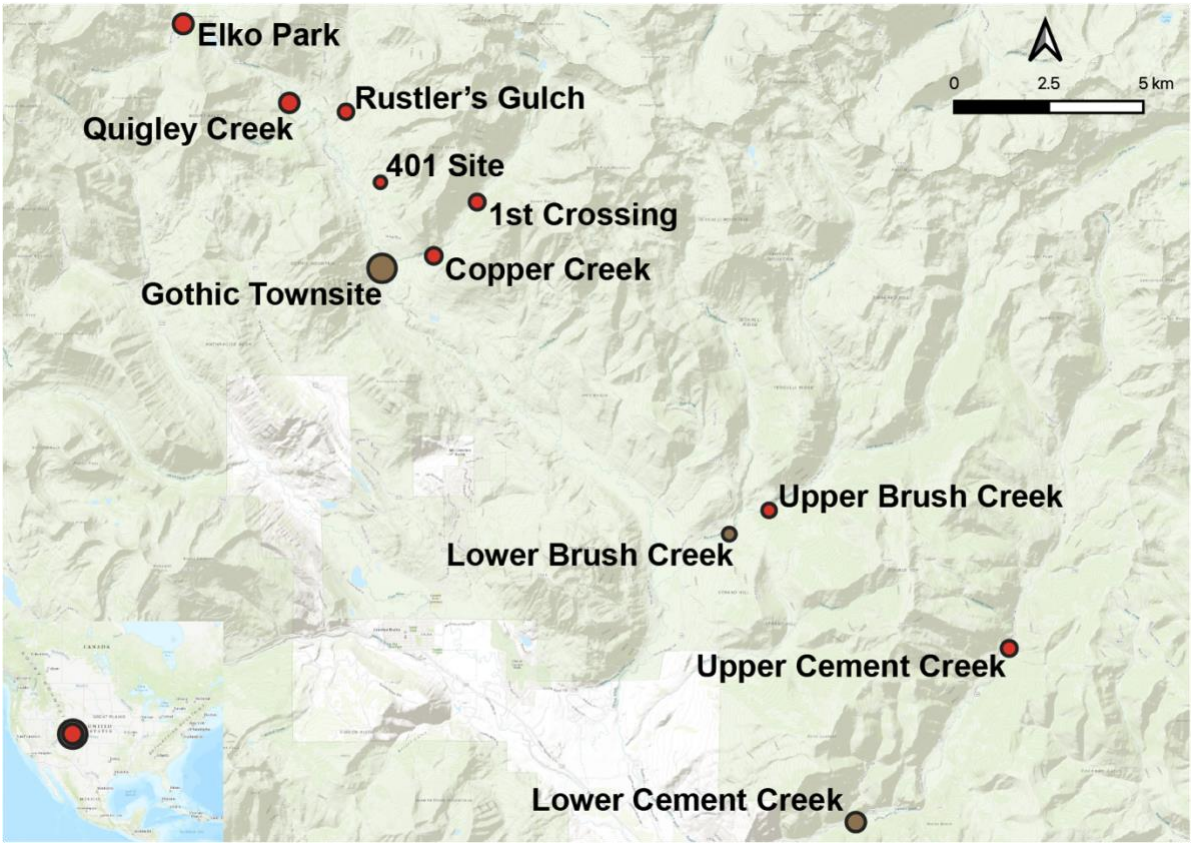

Figure 2.1: Map of the sites used in this study. Sites in red represent areas where *Thlaspi arvense* does not occur and sites in brown represent areas where *T. arvense* is present. The size of the circle represents the magnitude of sample sizes from each location.

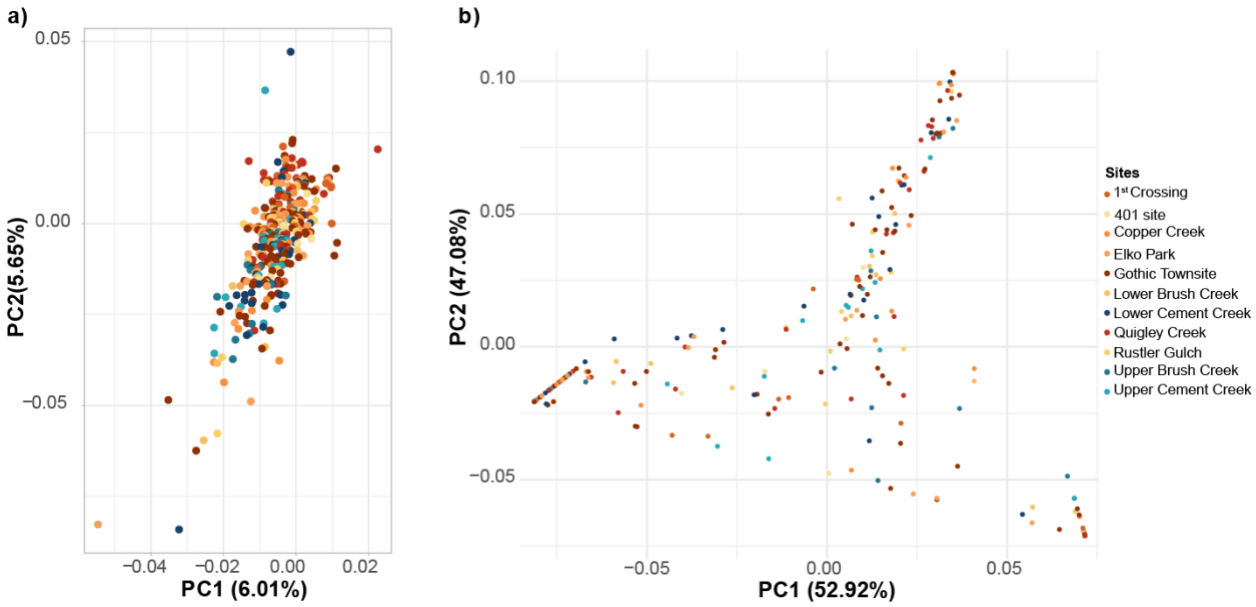

Figure 2.2: a) Genome wide Principal Component Analysis (axes 1 and axes 2) using neutral loci indicating a mixed population without distinct site-specific clusters. **b)** Archetypal analysis (axes 1 and axes 2) with k=3 indicating a mixed population without distinct site-specific clusters.

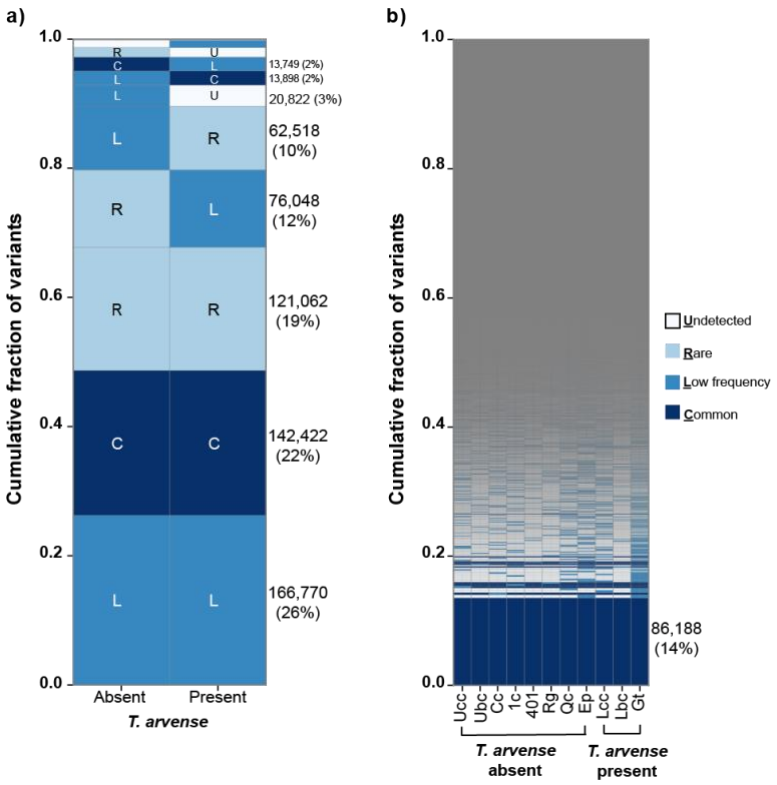

Figure 2.3: a) The relative abundance of genomic allele frequency variants of *P. macdunnoughii* in areas with *T. arvense* and areas without *T. arvense*. The codes depicted in the figure represents the frequency of the minor allele in the habitat. U-undetected: No alleles present, R-rare: <1% minor allele frequency (MAF), L-low frequency: 1%-5% MAF, C-common: >5% MAF. The percentages and their corresponding numbers indicate the number of variants in each class. Grey rows represent alleles whose contribution was non-significant to the differences between comparisons. **b)** Relative abundance of variants of *P. macdunnoughii* in our study area. Ucc: Upper Cement Creek, Ubc: Upper Brush Creek, Cc: Copper Creek, 1c: 1st Crossing, 401: 401 site, Rg: Rustler's Gulch, Qc: Quigley Creek, Ep: Elko Park, Lcc: Lower Cement Creek, Lbc: Lower Brush Creek, Gt: Gothic townsite.

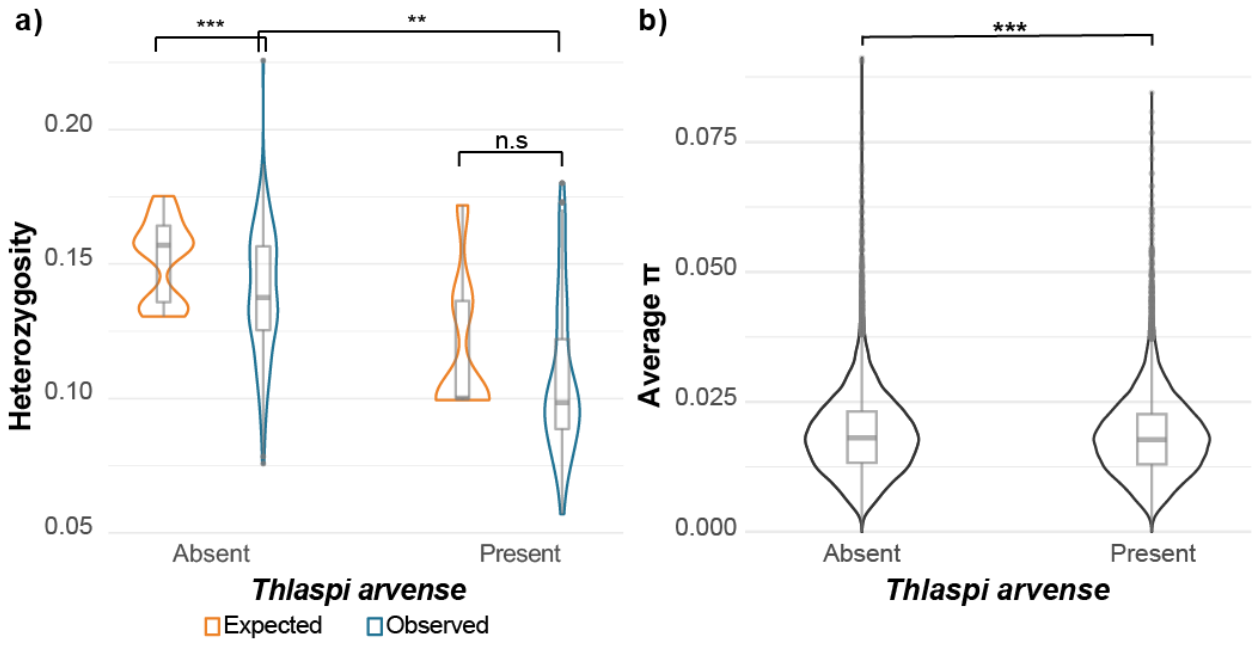

Figure 2.4: a) Genome wide variance in expected and observed heterozygosity of *P. macdunnoughii* in the ERV. ***P<0.001, **P<0.01, n.s.=not significant. Bartlett's test comparing expected and observed heterozygosity in areas without *T. arvense* (Bartlett's K^2=39.63, df=1, p<0.001), in areas without *T. arvense* (Bartlett's K^2 =1.2707, df=1, p=n. s) and observed heterozygosity between areas with and without *T. arvense* (Bartlett's K^2 =6.95, df=1, p<0.01). **b)** Genome-wide variance in nucleotide (π) diversity of *P. macdunnoughii* in areas without and with *T. arvense* (Bartlett's K^2= 9.6144, df=1, p=0.001).

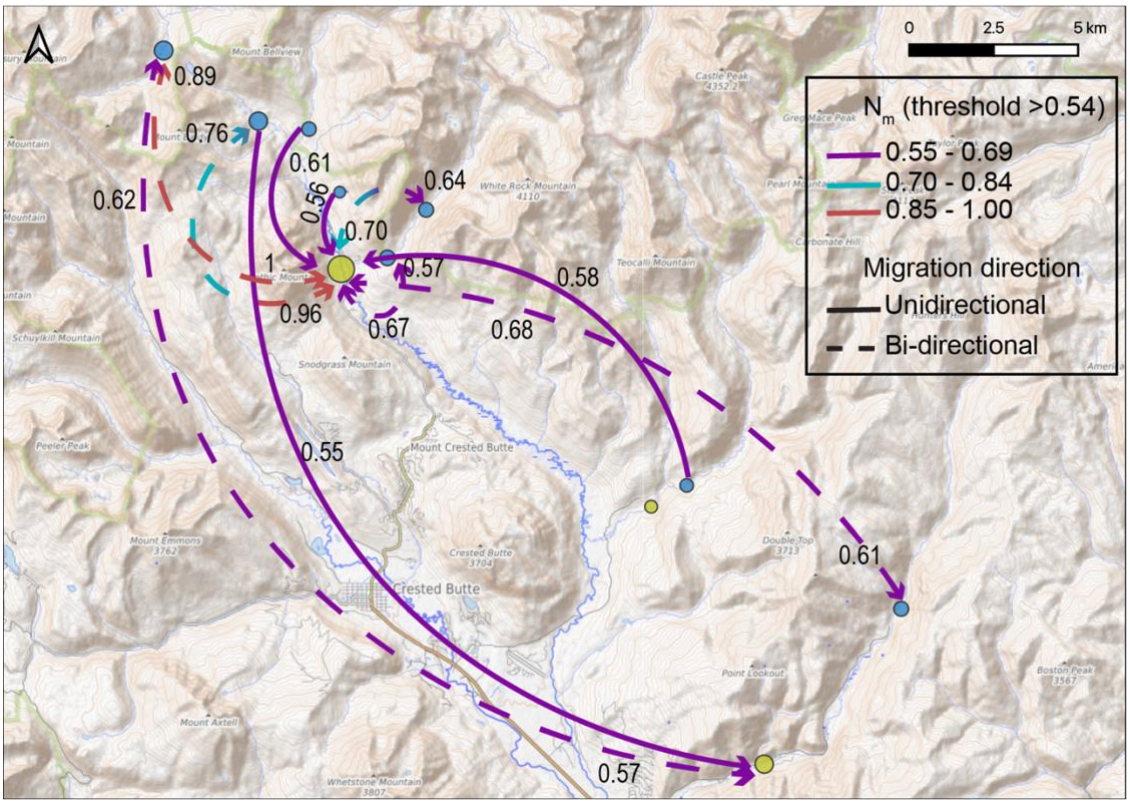

Figure 2.5: Relative migration rates of *P. macdunnoughii* among the 11 sites in the East River Valley estimated using divMigrate based on Nm (number of migrants/generations). The colors indicate migration levels, and the arrows indicate direction of migration. Solid lines represent unidirectional migration and dashed lines represent bi-directional migration. Migration routes are only shown for Nm > 0.55 and migration routes from areas with *T. arvense* to areas without *T. arvense* and vice-versa.

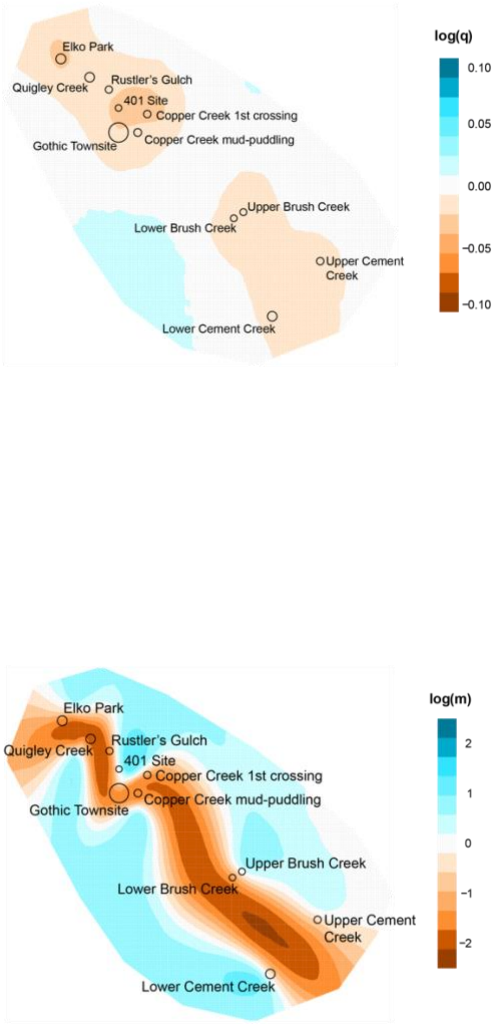

Figure 2.6: a) Estimated Effective Migration Surfaces and **b)** Estimated Effective Diversity Surfaces for *P. macdunnoughii* in the East River Valley. The migration and diversity rates, log(m) and log(q) represent gene flow/genetic diversity barriers and corridors in the habitat respectively. Each value corresponds to a 10-fold increase (blue) or decrease (orange/brown) in migration/genetic diversity compared to the null hypothesis of isolation by distance (white).

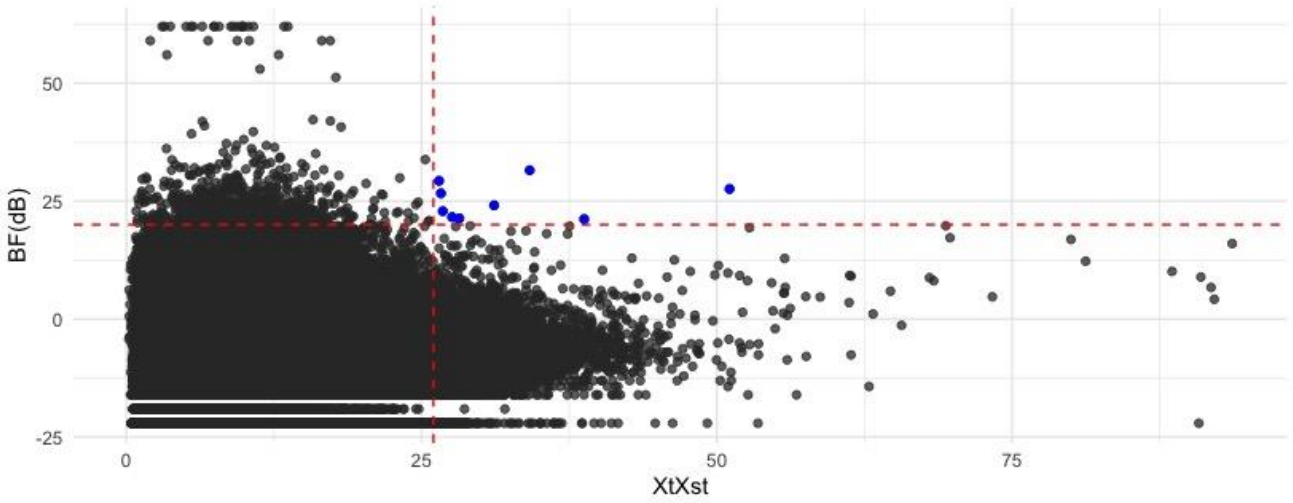

Figure 2.7: Pairwise comparison of median SNP XtX and BFmc values from 3 independent runs for adaptive differentiation associated with *T. arvense* presence in the habitat. The vertical dashed line represents the 1% POD (Pseudo Observed Dataset) significance threshold (XtX=26), and the horizontal dashed line represents the 20-dB threshold for BFmc. Blue dots represent the 9 loci of interest (outliers for both XtX and BFmc).

CHAPTER 3

DIFFERENTIAL GENE EXPRESSION UNDERLYING A NATIVE HERBIVORE'S OVIPOSITION PREFERENCE AND LARVAL PERFORMANCE ON A TOXIC INVASIVE PLANT.[3]

[3] Ravikanthachari N, Boggs CL. To be submitted to the Proceedings of the Royal Society B

Introduction:

Herbivorous insects and the plants they feed on are a classic example of co-evolution (Ehrlich & Raven, 1964; Thompson, 1999a, 1999b). Herbivorous insects have evolved to feed on a narrow range of hostplants by adapting to circumvent specialized plant defenses (Hardy & Otto, 2014; Jaenike, 1990; Joshi & Thompson, 1995). In holometabolous insects such as Lepidoptera, where the larval stages are relatively immobile, the success of the herbivore depends on the female choosing a suitable host that maximizes the survival of the resulting offspring (Berdegué et al., 1998; Bossart & Scriber, 1999; Ladner & Altizer, 2005). This results in strong selection for female oviposition preference. This idea, known as preference-performance hypothesis, postulates that the preference of a hostplant by the female is positively correlated with performance of the larvae on that hostplant (Jaenike, 1978; Valladares & Lawton, 1991). Female host choice is dependent on a multitude of factors including availability of suitable hosts, hostplant specialization, hostplant density and presence of natural enemies (Carrasco et al., 2015; García-Robledo & Horvitz, 2012; Larsson & Ekbom, 1995). Support for the preference-performance hypothesis ranges from none to strong correlation in various insect groups (Menacer et al., 2021; Rausher, 1979; Valladares & Lawton, 1991). A meta-analysis showed that oligophagous insects had a stronger correlation compared to polyphagous insects (Gripenberg et al., 2010). However, for specialized insects (those that feed on a single hostplant family), the evidence is mixed. One underlying hypothesis for this mixed evidence is that cue similarity between hosts and non-

hosts often leads to error prone oviposition in females (Steward & Boggs, 2020; Trowbridge & Todd, 2001).

The consequences of error prone oviposition can be outlined by the following scenarios (Wiklund, 1975). Error prone oviposition can be adaptive if the larvae can survive on the novel hostplant, thus leading to an increase in the diet breadth of the herbivore and adaptation to new hostplants (Janz et al., 1994; Stefanescu et al., 2012). In fact, error prone oviposition has often been cited as the fuel driving hostplant range in Lepidoptera and has been documented in many species (Janz et al., 1994; Nylin et al., 2000; Nylin & Janz, 1993, 2009). The success of the larvae on the novel host depends on underlying variation in detoxification enzymes, the degree of chemical similarity of the novel host to the hosts that the larvae have co-evolved with and the evolutionary history of host use in the insect species (Calla et al., 2017; Celorio-Mancera et al., 2016; Erbilgin et al., 2014; Futuyma & Agrawal, 2009).

Alternatively, error prone oviposition can lead to high fitness costs if the larvae do not possess behavioral or physiological mechanisms to incorporate the novel host plant, thus leading to higher mortality and maladaptation (Casagrande & Dacey, 2007; Keeler & Chew, 2008; Steward et al., 2019a; Steward & Boggs, 2020; Yoon & Read, 2016). Repeated error prone oviposition can potentially lead to the larvae evolving the ability to feed on the plant provided that new mutations or underlying standing genetic variation leads to evolution of key innovations in detoxification machinery (Nylin & Janz, 1993; Wheat et al., 2007).

In certain scenarios, known as evolutionary traps, abrupt or rapid environmental change can lead to instances where evolved reliable cues fail, resulting in repeated oviposition error (Keeler & Chew, 2008; Singer, 2021; Singer & Parmesan, 2019b; Steward & Boggs, 2020; Yoon & Read, 2016). This could occur either due to the female's inability to distinguish between hosts and non-hosts to avoid costly mistakes or the failure of larvae to incorporate the novel host due to behavior or physiological constraints (Casagrande & Dacey, 2007; Chew, 1977b; Huang et al., 1994; Keeler & Chew, 2008; Singer & Parmesan, 2019b; Steward et al., 2019a). Identifying the underlying mechanisms that leads to decoupling cue-response systems as well as preference-performance can shed light on the constraints of evolution in species interactions.

Although evolutionary traps have been documented in many systems, mechanistic processes underlying persistent evolutionary traps are not well understood (Hale et al., 2016; Robertson et al., 2013; Robertson & Chalfoun, 2016; Robertson & Hutto, 2006). Here we use a system of a specialized native herbivore that oviposits on a lethal, invasive plant and whose larvae have 100% mortality to dissect the underlying genetic mechanisms affecting adult female oviposition, and larval feeding on the plant. We employ a transcriptomic approach to quantify differential gene expression in order to address the hypotheses that a) females preferring the toxic lethal plant differ in their expression repertoire of sensory genes compared to those preferring the native plant and b) larvae that feed on the toxic plant exhibit transcriptomic signatures of impaired feeding and toxicity (upregulation of stress responses) compared to

those feeding on the native host plant. Our study adds to the extensive literature on preference-performance studies by elucidating the underlying genetic mechanisms that decouple female preference and larval performance.

Methods:

Study organisms

Pieris macdunnoughii Remington, 1954 (Pieridae; formerly *P. napi macdunnoughii*) (Chew & Watt, 2006) is an univoltine montane butterfly distributed in the Southern Rocky Mountains in North America. *Pieris macdunnoughii* is a specialist herbivore: the females lay eggs and the larvae feed on native Brassicaceae (Chew, 1977b, 1977a, 1980; Nakajima & Boggs, 2015). *Pieris macdunnoughii*, like other species in the *Pieris* species complex, has evolved detoxification enzymes to overcome the toxicity of glucosinolates (secondary metabolites) in Brassicaceae (Edger et al., 2015; Wheat et al., 2007). However, they perform poorly on species of mustards whose glucosinolates differ from those they have locally co-evolved with (Chew, 1977a).

Thlaspi arvense (L.) (Brassicaceae), a plant native to temperate Eurasia and has been introduced to the North America in the 1800s (Chew, 1977b). *Thlaspi arvense* was introduced to the Gunnison County between 1850s to 1970s and thus has been present for about 45 - 150 years (Chew, 1977b). *Thlaspi arvense* is an early successional plant. It colonizes disturbed soil and is known to occur up to 2900m (Best & McIntyre, 1975; Warwick et al., 2011).

The chemical profile of *T*. *arvense* is comprised mainly of the aliphatic glucosinolate sinigrin, compared to native mustards that contain both aliphatic and aromatic glucosinolates (Rodman & Chew, 1980). The glucosinolate sinigrin acts as an oviposition attractant and hence females choose to lay eggs on *T. arvense* as it contains higher sinigrin concentration compared to native plants (Steward & Boggs, 2020). *Pieris macdunnoughii* larvae are able to metabolize low concentrations of sinigrin present in the native plants but the higher concentration present in *T. arvense* is toxic to the larvae, thus indicating that the differences in sinigrin concentration drives the maladaptation (Steward et al., 2019a). *Thlaspi arvense* is completely lethal to the butterfly and none of the larvae survive past pupation in the lab or field (Chew, 1977b; Steward et al., 2019a). Co-occurrence of *T. arvense* with other native mustards has been estimated to cause a fitness loss to females of 3% due to larval survival, habitat occupancy by the females and the distribution and abundance of *T. arvense* compared to the native plants in the habitat (Nakajima et al., 2013; Nakajima & Boggs, 2015). Therefore, there is strong selection on females of *P. macdunnoughii* to avoid laying eggs on the plant and for escape from the evolutionary trap.

Oviposition choice experiment

Gravid females for the oviposition choice experiment were collected in the East River Valley, Gunnison County, Colorado (38° 57' 29.1096" N, 106° 59' 21.4404" W) in 2019 using an aerial net. In the lab, the females were fed twice a day with 25% honey-water solution. Females were housed in plastic cages in an

environmental chamber at 27°C during the day and at 18°C at night on an 18:6 L:D cycle. The females were provided with one whole plant each of *T. arvense* and *Cardamine cordifolia* A. Gray (Brassicaceae), a primary native hostplant. The larval host plants were matched by approximate leaf area and by plant phenology (pre-flowering stage). The females were allowed to lay eggs on the host plants and the eggs from each plant were counted every morning. Preference for a hostplant was quantified as follows: preference for *T. arvense* if >80% of eggs laid on *T. arvense*; preference for *C. cordifolia* If >80% of eggs laid on *C. cordifolia* and equal preference if < 80% eggs laid on either plant. The females were removed from the choice experiment once they had laid a total of 50 eggs and their forelegs, antennae and head were stored in RNAlater at -20°C.

Larval feeding assays

Eggs laid by the females in the oviposition choice trails were used for the larval feeding assays. The larvae from hatched eggs were reared in an environmental chamber under the same conditions used during female oviposition experiment. Larvae were reared on young leaves of *Raphanus sativus* L. (Brassicaceae) until they reached third instar. Five 3rd instar larvae from each female were provided either with a whole plant of *C. cordifolia* or *T. arvense* after 12 hours of starvation. Larvae were allowed to feed for 24 hours. After 24 hours, the larvae were visually examined to confirm that their midguts were completely filled with plant material. The larvae were then dissected in PBS solution, degutted, plant material removed, and their mouth parts and guts were stored in RNAlater.

RNA extraction and sequencing

Total RNA was extracted from 10 females (head, antennae, and forelegs) each preferring *T. arvense* and *C. cordifolia* and from 10 third instar larvae (gut and mouthparts) each fed on *C. cordifolia* and *T. arvense* using a QIAGEN RNeasy kit following manufacturer's protocol. Total RNA from all samples were sent to MedGenome for cDNA library preparation and RNA sequencing. cDNA library preparation was carried out using Illumina TruSeq stranded mRNA kit followed by 20M paired reads (40M total) sequencing on a Novaseq S4 platform.

Read mapping

Demultiplexed raw Illumina reads were checked for quality using FastQC (Andrews, 2010). Reads were trimmed using Trimmomatic v. 0.39 (Bolger et al., 2014) with the following options: Sliding window: 4:20, minimum length: 25, and adapter sequences were cleaned using the option TruSeq3-PE in Trimmomatic. Cleaned transcripts were assembled using genome guided STAR assembly (Dobin et al., 2013; Dobin & Gingeras, 2015) using the P. macdunnoughii genome (Steward et al., 2021). We first generated genome indices and the genome directory, followed by mapping reads to the genome with default options. 94% of the reads were correctly mapped to the genome. The resulting BAM files were sorted by genome coordinates using SAMtools (Li et al., 2009a) sort and indexed using SAMtools index. Mapped reads were then quantified using featureCounts (useMetaFeatures geneid) (Liao et al., 2014) and ran through Rsubread (Liao et al., 2019) in the R statistical environment v. 4.0 (R

Development Core Team, 2008). We ran featureCounts separately on adult female and larval mapped reads to obtain their respective expression profiles.

Differential gene expression analysis

Differential gene expression (DGE) analyses for adult females and larvae were carried out separately in edgeR (M. D. Robinson et al., 2010). We used CPM-TMM log2 transformation for further filtering our dataset. The raw counts were filtered based on abundance and then normalized using count-per-million (CPM) to account for differences in library sizes among samples in edgeR and trimmed mean of M values (TMM) for cross-sample normalization using the option "calcNormFactors" in edgeR. Genes that were lowly expressed were filtered using the HTSFilter followed by differential gene expression analysis using the Fisher's exact test approach in edgeR. The differentially expressed genes were further filtered using a false discovery rate of 1% using the BH correction method. We set a cut-off of a minimal fold change (FC) of three between treatments and an FDR of p-value <1e-3 for assessing significant differential expression. The cut-off values for FC and FDR were set based on prior studies to increase the signal: noise and to reduce non hostplant specific patterns (Breeschoten et al., 2019; Breeschoten, Schranz, et al., 2022). We used the hclust function in R package stats for hierarchical clustering of significantly differentially expressed transcripts with a k value of 5 based on k-means clustering. We additionally visualized differentially expressed genes using a heatmap using the pheatmap package in R (Kolde & Kolde, 2018).

Gene set enrichment

We generated gene ontology (GO) annotation of the *P. macdunnoughii* genome using eggNOG-Mapper (Huerta-Cepas et al., 2017) and aligned the annotated genes to the Pfam database (Finn et al., 2014). The GO annotation consisted of ~12000 genes that were used as input for topGO (Alexa & Rahnenführer, 2009) to quantify gene set enrichment in the differentially expressed genes. We calculated gene set enrichment separately for each of the 5 distinct hierarchical gene clusters. We used a weighted two-tailed Fisher's test to identify enriched GO terms based on molecular function (MF), biological processes (BP), and cellular components (CC). The enriched GO terms were then run through REVIGO (Supek et al., 2011) to cluster the terms and to reduce redundancy by identifying similarity between the GO terms. The GO terms were additionally grouped into clusters by assigning the terms to their parent molecular function (MF) description based on Pfam annotation in order to classify MF that were enriched in each treatment. We ran a paired t-test to assess significance in gene expression between the treatments in the larvae.

Results:

Variation in female oviposition preference

We tested 56 wild caught female *Pieris macdunnoughii* collected from the East River valley. 14 females preferred *Cardamine cordifolia*, and 10 females preferred *Thlaspi arvense*. Nine females had equal preference for both plants

and the preference of 25 females did meet the threshold of 50 eggs to quantify preference.

No differential gene expression underlying female oviposition preference

Gene expression profiles of female *P. macdunnoughii* preferring *T. arvense* and *C. cordifolia* were similar. We did not identify any genes that were differentially expressed (DE) between them. Our analysis consisting of females who laid 100% of their eggs on either of the plant (n=3/group, total 6 females) also did not identify any DE genes.

Hostplant-specific changes in differential expression of genes underlying larval feeding

We found differences between gene expression profiles of larvae that fed on *T. arvense* and *C. cordifolia*. Hierarchical clustering of individuals showed that replicates of larvae of each treatment were similar to each other compared to those from the other treatments (fig. 3.1a). However, one of the larvae that fed on *C. cordifolia* clustered separately from all other larvae and showed distinct expression profile. The hierarchical clustering of transcripts revealed 196 genes (FC >3, FDR <0.05) that were differentially expressed between larvae that fed on *C. cordifolia* compared to *T. arvense*. These transcripts split into 5 distinct clusters (fig. 3.1a). Clusters 1 (n=76 genes; table 3.1; paired t-test, $p<0.001$) and 2 (n=20 genes; table 3.1; paired t-test, $p<0.001$) were upregulated in larvae that fed on *C. cordifolia*. Cluster 3 (n=33; table 3.1; paired t-test, $p<0.001$) had similar expression patterns in all but one larva that fed on *C. cordifolia*, where all the

genes were highly upregulated in that particular larvae. Clusters 4 (n=32; table 3.1; paired t-test, $p<0.001$) and 5 (n=35; table 3.1; paired t-test, $p<0.001$) were upregulated in larvae that fed on *T. arvense* (fig. 3.1b).

Gene set enrichment analysis reveals host-plant specific effects on critical larval traits

Cluster 1 had GO terms related to metabolic processes and cellular transport overrepresented in the biological processes (BP) category in larvae feeding on *C. cordifolia*. In the molecular function (MF) ontology category, cluster 1 had 19 GO terms primarily related to transporter activity overrepresented. Cluster 2, which was upregulated in larvae feeding on *C. cordifolia* had GO terms related to development and immune cell proliferation overrepresented in the BP category and GO terms related to binding and transferase activity overrepresented in the MF category (fig. 3.1b, table 3.1).

Cluster 3 showed GO term overrepresentation in only one larva in the *C. cordifolia* treatment. GO:0042335 cuticle development and GO:0019233 sensory perception of pain was overrepresented in that larva (fig. 3.1b, table 3.1).

Transcripts in cluster 4 in larvae feeding on *T. arvense* had GO terms related to immune responses and responses to stimuli overrepresented in the BP category. Transcripts in cluster 5 which were also upregulated in the *T. arvense* treatment larvae showed GO terms related to lipid metabolism, cell death and ageing overrepresented in the BP category (fig 3.1b, table 3.1).

Molecular functions underlying various processes differentially expressed between diet treatments

We found significantly higher expression in catalytic activity (fig. 3.2; table 3.2; paired t-test, p=0.005), cellular component (fig. 3.2; table 3.2; paired t-test, p=0.015), localization (fig. 3.2; table 3.2; paired t-test, p=0.02) and transporter activity (fig. 3.2; table 3.2; paired t-test, p<0.001) in larvae feeding on *C. cordifolia* (fig. 3.2, table 3.2) compared to *T. arvense*. Larvae feeding on *T. arvense* had higher expression in processes related to ATP-dependent activity (fig. 3.2; table 3.2; paired t-test, p=0.02), metabolic process (fig. 3.2; table 3.2; paired t-test, p=0.012), and response to stimulus (fig. 3.2; table 3.2; paired t-test, p=0.005) compared to those feeding on *C. cordifolia*. Molecular functions involved in antioxidant activity, binding, cellular processes and transport, molecular carrier activity, oxidoreductase activity and structural molecular activity had similar expression across the two treatments.

Discussion:

We compared gene expression patterns of adult females of *Pieris macdunnoughii* when laying eggs and third instar larvae when feeding on the toxic, invasive plant *Thlaspi arvense* with that of *Cardamine cordifolia*, a native host plant. Although we saw variation in oviposition choice by the females in our assays, we did not find any differential gene expression between the females preferring different plant species. We however found 196 transcripts that were differentially expressed in the larvae in the two treatments. We found host-plant

specific gene expression patterns in the larvae and identified underlying gene ontologies that were enriched when feeding on the native plant or on *T. arvense*. Larvae feeding on *C. cordifolia* showed upregulation of metabolic processes, development, and immune cell proliferation, suggesting favorable hostplant feeding, whereas larvae feeding on *T. arvense* showed upregulation of immune processes, responses to stimuli, ageing, fatty acid metabolism and cell death, indicating signs of impaired feeding. Overall, our results dissect the underlying mechanisms of the evolutionary trap and identify the mechanisms decoupling preference-performance in *P. macdunnoughii* in the presence of the invasive plant, *T. arvense*. Our work also provides new insights on herbivore responses to novel resources by combining both adult and larval responses and identifying the ontogenetic stage where adaptation was likely to occur.

Female responses to *Thlaspi arvense*.

Pieris macdunnoughii has been experiencing *T. arvense* for at least 45 generations (since the evolutionary trap was identified) and possibly for 150+ generations (since *T. arvense* was first introduced to the Gunnison Valley) (Best & McIntyre, 1975; Chew, 1977b). Given that fitness costs of laying eggs on *T. arvense* is high, evolution of female preference to avoid laying on the plant would be expected. Although Lepidoptera females use a combination of olfactory, visual, and gustatory cues to lay eggs on suitable hostplants, the ultimate step involves gustatory recognition (Pivnick et al., 1994; Steward & Boggs, 2020). Gustatory receptors are localized in the forelegs of the females and olfactory and visual receptors are located on the antennae and eyes respectively (McIndoo,

1929; Xu, 2020). Thus, any differences in perception should be identifiable in these organs. However, we did not find any differences in gene expression in the antennae, head, and forelegs of the females when choosing between the two hostplants. This suggests that females are unable to differentiate between the two hostplants and therefore, error-prone oviposition on *T*. *arvense* is the underlying cause in the persistence of the evolutionary trap. *Thlaspi arvense* and *C. cordifolia* both contain sinigrin, an aliphatic glucosinolate that is an oviposition attractant (Rodman & Chew, 1980; Steward et al., 2019a). Sinigrin concentration is, however, magnitudes higher in *T. arvense* compared to the other native hostplants and therefore acts as a stronger oviposition attractant compared to the native plants (Steward et al., 2019a; Steward & Boggs, 2020). Although females can differentiate between different oviposition attractants through differential gene expression, the gustatory receptors might not be fine-tuned to differentiate between different concentrations of the same compound (Orsucci et al., 2018; Ozaki et al., 2011). Thus, sinigrin in *T. arvense* and the native plants might be presenting as the same cue, albeit at a presumed higher fitness, thus driving females to oviposit on *T. arvense*.

Larval responses to *Thlaspi arvense*.

We found 196 transcripts that were differentially expressed in larvae between the two treatments. Hierarchical clustering revealed that larvae in the same treatment were similar to each other compared to the other treatment in their expression profiles, thus suggesting hostplant specific effects on larval gene expression. Of the five distinct gene clusters, two showed upregulation in larvae

feeding on *C. cordifolia* and two showed upregulation in larvae feeding on *T. arvense*.

Transcripts in cluster 1 (table S3.1) showed upregulation in GO terms related to metabolism including carbohydrate metabolism (GO:0005975; GO:0005996; GO:0008643; GO:0034219), vitamin metabolism (GO:0006766; GO:0009812; GO:0035461), organic acid metabolism (GO:0006082; GO:0015882; GO:0019852), uronic acid metabolism (GO:0006063) among others. Cluster 1 also showed upregulation of several growth-related GO terms including imaginal disc pattern formation (GO:0007447), wing disc pattern formation (GO:0035222), neuron recognition (GO:0008038; GO:0035082), respiratory tube (GO:0030323), and lung development (GO:0030324), among others. Several GO terms related to biosynthetic process were also upregulated including folic acid biosynthesis (GO:0009396) and pteridine biosynthesis (GO:0042559). GO terms related to adaptation, such as general adaptation syndrome (GO:0051866), and adaptation syndrome to behavioral process (GO:0051867) were also upregulated. Processes mentioned above are hallmarks of adaptive larval performance and are correlated with higher larval survival and pupation. Apart from these, GO terms related to lipid storage including negative regulation of fatty acid metabolism and lipid export from cell, which are positive indicators of high-quality resource were also upregulated (Boggs, 1981; Boggs & Freeman, 2005; D. A. Hahn, 2005; Liu et al., 2009).

Transcripts in cluster 2 (table S3.2) showed upregulation in GO terms related to development including pupation (GO:0035074; GO:0035209), positive

regulation of epidermis development (GO:0045682; GO:0045684), intestinal epithelial cell development (GO:0060575; GO:0060576), response to ecdysone (GO:0071390) and reproduction (GO:0000003; GO:0000909; GO:0048608; GO:0048806; GO:0001893). GO terms related to immune cell proliferation including T cell differentiation (GO:0033078; GO:0033082), Natural Killer T cell differentiation (GO:0001779; GO:0001865; GO:0030101), cell proliferation involved in imaginal disc-derived wing morphogenesis (GO:0090256) and negative regulation of leukocyte proliferation (GO:0070661) was also enriched. Plant detoxification related GO terms were also enriched such as UDPG metabolic process (GO:0006011), response to sterol (GO:0036314; GO:0036315) and positive regulation of response to oxidative stress (GO:1902884). Sex organ development (gonads and ovaries) usually develop during the third instar in several Pieridae and can be visually seen in the dorsal surface of the larvae. Sex organ development is positively correlated with hostplant quality in several insects, including Lepidoptera (Awmack & Leather, 2002; Boggs, 1981, 2009; Clissold & Simpson, 2015; Ishihara & Ohgushi, 2006; Landolt & Phillips, 1997). Immune cell proliferation is also known to be upregulated in other phytophagous insects when feeding on suitable hostplants (Diamond & Kingsolver, 2010; Klemola et al., 2007). Detoxification enzymes such as UDPGT are involved in plant metabolite detoxification where UGT enzymes mediate the transfer of glycosyl residues from nucleotide sugars to acceptor molecules that are then converted to lipophilic substances which are excreted from the body (Tephly & Burchell, 1990).

Transcripts in cluster 3 (table S3.3) showed upregulation in only one larva that fed on C. cordifolia. Based on the GO terms enriched, such as cuticle development and sensory perception to pain, this individual is likely an outlier and the expression profiles patterns are either due to injury prior to sample processing or other unaccounted damage.

Cluster 4 (table S3.4) showed upregulation in larvae that fed on *T. arvense*. GO terms related to immune response, toxin metabolic process and cellular damage were upregulated. GO terms that indicated immune response included regulation of immune response (GO:0002682; GO:0002684; GO:0002697), antimicrobial humoral response (GO:0019730), response to bacterium (GO:0009607; GO:0009617; GO:0042742; GO:0050830; GO:0002833; GO:0031349) and toll signaling pathway (GO:0008063; GO:0008592; GO:0045752). The above immune responses indicate immune responses to fungal and/or bacterial pathogens in the midgut. For example, Toll signaling pathway is upregulated in *Drosophila* in response to bacterial infection (Anderson, 2000). Plant defenses are known to interact with gut microbiome and can alter the composition from mutualistic bacteria to pathogenic bacteria, especially when larvae feed on low quality or non-host species by initiating leaky guy syndrome (Hammer & Bowers, 2015; Mason et al., 2019). Coupled with immune responses, GO terms related to toxin metabolism including toxin catabolic process (GO:0009407; GO:0009636) and xenobiotic catabolic process were enriched. Additionally, GO terms related to cellular damage such as cell killing (GO:0001906; GO:0031341; GO:0031343) and cytolysis (GO:0019835;

GO:0042268) were enriched. Overall, gene expression profile of cluster 4 indicates plant defense mediated pathogenic changes in microbiome that impair feeding and growth on *T. arvense*.

Cluster 5 transcripts (table S3.5) showed upregulation in *T. arvense* feeding larvae in GO terms broadly related to lipid metabolism, response to starvation, aging and cell death. GO terms related to lipid metabolism that were enriched included fatty acid biosynthesis pathway (GO:0006629; GO:0006633; GO:0001676; GO:0006636; GO:0008610; GO:0016042; GO:0033559; GO:1901571), phospholipid biosynthesis pathway (GO:0006644), phosphatidylcholine catabolic process (GO:0034638; GO:0036151; GO:0036152) and eicosanoid metabolic process (GO:0046456). Increased lipid metabolism coupled with GO terms related to response to starvation (GO:0042594) indicate that the larvae are unable to metabolize *T. arvense* and thus starving to death. *P. macdunnoughii* larvae have been shown to initiate feeding later on *T. arvense* and pre and post ingestive defenses are known to deter feeding, thus increasing starvation and mortality (Steward et al., 2019a). In holometabolous insects, adult lipid content is dependent on larval lipid storage as most of the lipids are derived during larval development (Boggs, 1981; Boggs & Freeman, 2005). Increased lipid catabolism in larval stages is known to impact successful pupal eclosion, adult survival, reproduction, and various life history traits (Boggs, 1981). Low lipid reserves in larvae is positively correlated with higher mortality.

Additionally, our results of Pfam function enrichment between the two treatments indicated higher expression of transporter activity, localization, and catalytic activity in larvae feeding on *C. cardamine*, further providing support that larvae feeding on *C. cardamine* have higher success compared to those feeding on *T. arvense* where functions related to response to stimulus (immune response, xenobiotic compounds), metabolic process and ATP - dependent activity were higher in their expression.

Overall, the larval gene expression profiles show stark contrasts among the different clusters and molecular functions, where upregulation of transcripts in clusters 1 and 2 is indicative of a high-quality host plant (*C. cordifolia*) and upregulation of clusters 4 and 5 indicates a poor hostplant (*T. arvense*) that leads to the death of the larvae.

Implications for preference-performance hypothesis and the outcome of evolutionary trap.

Our results demonstrate that females' lack the ability to differentiate between the host plants and larvae show impaired signs of feeding on *T. arvense*. The fitness costs related with laying eggs on *T. arvense* is dependent on the fine-grained structure of *T. arvense* in the habitat and its proximity to native host plants (Nakajima et al., 2013; Nakajima & Boggs, 2015). *Thlaspi arvense* is also expected to increase in abundance due to increased droughts, warmer temperatures, and anthropogenic disturbances. This suggests that error-prone oviposition on T. arvense by *P. macdunnoughii* will increase due to an

increase in *T. arvense* abundance compared to the native host plants that prefer wet soils and colder climates. Analysis of signatures of selection in *P. macdunnoughii* in response to *T. arvense* indicated that local adaptation can occur in the face of high gene flow and low genetic variation (Ravikanthachari et al., unpublished). These results combined with the results from this study indicate that selection would likely act on larval ability to feed on *T. arvense* rather than females' ability to avoid laying eggs on the plant. Several cases where repeated error-prone oviposition by females have resulted in larvae incorporating the novel host plant has been documented, including in related species *Pieris oleracea* and its interaction with the invasive mustard *Alliaria petiolata* (Haribal et al., 2001; Huang et al., 1994). Larvae that can survive longer on *T. arvense* could potentially move to native hostplants in the later instars, resulting in rescue from the evolutionary trap.

Additionally, our analysis also highlighted gut microbiome changes in response to *T. arvense*. The role of gut microbiome in adaptation to new hosts has been elucidated in several species and might play a complementary role in *P. macdunnoughii* larvae adapting to *T. arvense* in the future (Blankenchip et al., 2018; Ge et al., 2021).

Conclusions:

Our results are among the few studies that have quantified both preference and performance at the level of gene expression. Most existing research has addressed only one component (either preference or performance)

at the level of gene expression, thus making it difficult to draw broader conclusions about the underlying mechanisms. Our results add to the decades of ecological studies that have addressed preference-performance in phytophagous insects by elucidating the genetic mechanisms underlying decoupling of preference-performance mismatches. Our results also suggest a role for the gut microbiome in response to novel hosts. This work provides a base for future work on hostplant-insect interactions to test the role of larval gut microbiome changes in mediating plant-insect interactions on novel hosts. This work combined with the extensive research on the *Pieris*-*Thlaspi* evolutionary trap system serves as a model to understand the continued effects and predict the outcomes of native species interactions with novel resources.

Acknowledgements:

We thank Malia Olson, Hannah Walton, Emma Wagner, Ojaswee Shrestha, and Skyler McDaniel for assisting with animal rearing. We thank Sushant Potdar for providing advice on adult antennae RNA extraction. We thank Ward B Watt, Joseph M Quattro, Brian Hollis, Peter Andolfatto and Minoti Asher for proving comments on earlier versions of the manuscript. This work was supported by the University of South Carolina Arts & Sciences to CLB and, the Rocky Mountain Biological Laboratory (RMBL) graduate student fellowship to NR. Field research permits were arranged by RMBL.

Table 3.1: Paired t-test comparing gene expression (log (CPM)) between *C. cordifolia* and *T. arvense* among different clusters.

Gene cluster	**y (log)**	**group1**	**group2**	**n (*C. cordifolia*)**	**n (*T. arvense*)**	**statistic**	**df**	**p**	**p.adj**
1	CPM	*C. cordifolia*	*T. arvense*	340	340	-10.920	636.194	<0.001	**<0.001**
2	CPM	*C. cordifolia*	*T. arvense*	190	190	7.354	313.047	<0.001	**<0.001**
3	CPM	*C. cordifolia*	*T. arvense*	260	260	-9.118	459.432	<0.001	**<0.001**
4	CPM	*C. cordifolia*	*T. arvense*	670	670	13.936	1292.493	<0.001	**<0.001**
5	CPM	*C. cordifolia*	*T. arvense*	360	360	3.522	575.110	<0.001	**<0.001**

Table 3.2: Paired t-test comparing gene expression (log (CPM)) between *C. cordifolia* and *T. arvense* among different Pfam molecular functions.

Molecular function	**y (log)**	**group1**	**group2**	**n (*C. cordifolia*)**	**n (*T. arvense*)**	**Statistic**	**df**	**p**	**p.adj**
antioxidant activity	CPM	*C. cordifolia*	*T. arvense*	10	10	1.473	10.516	0.17	0.238
ATP-dependent activity	CPM	*C. cordifolia*	*T. arvense*	10	10	-3.117	10.975	0.010	**0.020**
binding	CPM	*C. cordifolia*	*T. arvense*	330	330	0.351	645.554	0.726	0.726
catalytic activity	CPM	*C. cordifolia*	*T. arvense*	440	440	3.360	860.756	<0.001	**0.005**
cellular anatomical comp.	CPM	*C. cordifolia*	*T. arvense*	10	10	3.280	14.301	0.005	**0.015**
cellular process	CPM	*C. cordifolia*	*T. arvense*	110	110	0.593	214.943	0.554	0.597
cellular transport	CPM	*C. cordifolia*	*T. arvense*	10	10	1.046	10.001	0.323	0.377

localization	CPM	*C. cordifolia*	*T. arvense*	50	50	2.701	91.200	0.008	**0.019**
metabolic process	CPM	*C. cordifolia*	*T. arvense*	60	60	-2.971	108.491	0.004	**0.013**
molecular carrier activity	CPM	*C. cordifolia*	*T. arvense*	30	30	1.636	55.971	0.108	0.168
oxidoreductase activity	CPM	*C. cordifolia*	*T. arvense*	20	20	-2.171	34.065	0.037	0.065
response to stimulus	CPM	*C. cordifolia*	*T. arvense*	80	80	-3.332	158	0.001	**0.005**
structural molecular activity	CPM	*C. cordifolia*	*T. arvense*	30	30	-1.113	54.125	0.271	0.345
transporter activity	CPM	*C. cordifolia*	*T. arvense*	210	210	8.335	389.310	<0.001	**<0.001**

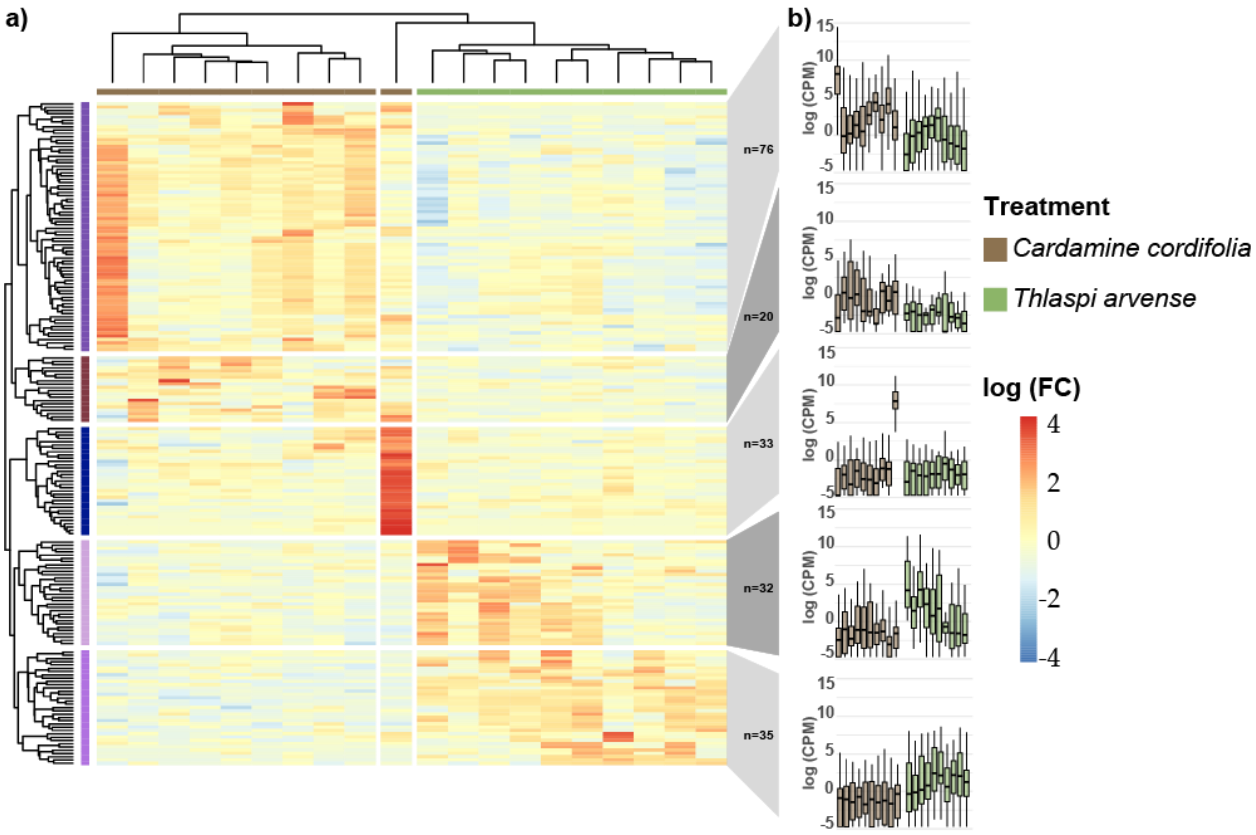

Figure 3.1: a) Heatmap showing log (FC) in larvae in response to feeding on *C. cordifolia* (brown) and *T. arvense* (green). Individuals (columns) and transcripts (rows) are arranged by hierarchical clustering of expression profiles. b) Individual gene expression profile corresponding to each cluster of gene expression represented in panel a.

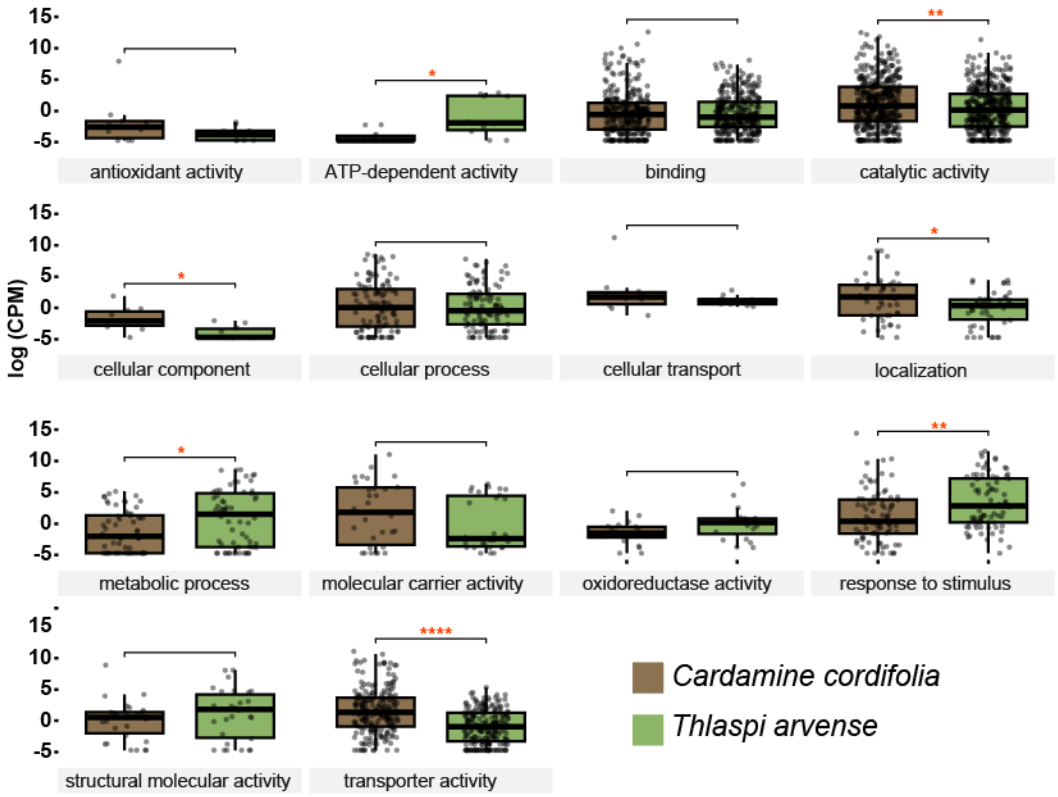

Figure 3.2: Gene expression of various molecular functions grouped based on Pfam database annotation between larvae feeding on *T. arvense* and *C. cordifolia*.

CHAPTER 4

GENE FAMILY EVOLUTION IN BRASSICACEOUS-FEEDING INSECTS: IMPLICATIONS FOR ADAPTATION AND HOST PLANT RANGE.[4]

[4] Ravikanthachari N, Boggs CL. To be submitted to Genome Biology and Evolution

Introduction:

Co-evolutionary interaction between plants and their herbivores is often proposed as one of the major drivers of eukaryotic biodiversity on earth (Ehrlich & Raven, 1964; Futuyma & Agrawal, 2009; Wiens et al., 2015). The co-evolutionary arms race is a product of ongoing key adaptive innovations in chemical and physical defense in plants and behavioral and chemical adaptation to circumvent those defenses in herbivores (M. Berenbaum, 1983; M. R. Berenbaum et al., 1986; Edger et al., 2015; Wheat et al., 2007). As a result, many herbivores have evolved specific mechanisms to detoxify a narrow range of hostplant chemicals and thus feed and develop on them (Hardy & Otto, 2014; Jaenike, 1990; Joshi & Thompson, 1995). The hostplant range of an herbivore is thus dependent on multiple factors including the degree similarity of chemical defenses among host plants, evolutionary history of hostplant use, genetic variation in detoxification enzymes and diversity of sensory repertoire in herbivores (Awmack & Leather, 2002; Bernays & Graham, 1988; Gripenberg et al., 2010).

Insect detoxification systems involve three distinct phases, which help in converting toxic plant chemicals to inert substances that are then excreted (Heidel-Fischer & Vogel, 2015; Kant et al., 2015; Rane et al., 2019). Phase I detoxification consists of cytochrome p450 monooxygenases (P450s or CYPs) and carboxylesterases (COesterase) which introduce reactive and polar groups into plant toxic compounds by reduction or hydrolysis (Dermauw et al., 2020; Feyereisen, 1999; Ranson et al., 2002). Phase II detoxification enzymes which

include Glutathione-S-Transferases (GSTs), UDP-glucosyltransferase (UDPGTs or UGTs) and Sulfotransferases add either glutathione, glucosyl or sulfate groups respectively to the byproducts from phase I or to toxic compounds themselves to increase their hydrophilicity (Cermak, 2008). Finally, enzymes in phase III such as ABC (ATP-binding cassettes) transporters that export these bound toxins to the extracellular matrix for excretion (Dermauw & Van Leeuwen, 2014).

Polyphagous herbivores (those that feed on a wide range of plant families with diverse chemical composition) and specialist herbivores (those that are specialized to feed on one plant family with a particular class of chemical profile) have different mechanisms by which they detoxify plant compounds. Polyphagous herbivores are known to have higher gene expression and increased gene copies of phase I and phase II general detoxification enzymes whereas specialist herbivores have evolved specific detoxification enzymes in addition to the general detoxification machinery (Calla et al., 2017; Edger et al., 2015; Wheat et al., 2007).

The interaction between the plant family Brassicaceae (mustard family) and its herbivores has been intensively studied in order to dissect the mechanisms underlying detoxification of specialized toxic compounds (Edger et al., 2015; Ratzka et al., 2002; Wheat et al., 2007). Brassicaceae have evolved glucosinolates (GLS), a class of secondary compounds that are stored separately from myrosinases (Wittstock & Halkier, 2002). Upon damage by herbivores, glucosinolates and myrosinases interact, followed by the hydrolysis of glucosinolates into highly toxic isothiocyanates (Halkier & Gershenzon, 2006).

Brassicaceae are well defended as most herbivores are unable to process isothiocyanates. However, specialized herbivores including some species of Lepidoptera such as *Pieris sps.* (Pieridae) and *Plutella xylostella* (Plutellidae) have evolved mechanisms to redirect the GLS-myrosinase reaction to instead form nitriles that are less toxic (Heidel-Fischer & Vogel, 2015). *Pieris sps.* have evolved nitrile specifier protein (NSP), which is expressed in the midgut to detoxify GLS (Chew, 1977a, 1980; Kuchernig et al., 2012). NSPs belong to the family of Insect allergen repeat, which is present in all insects and whose function is largely unknown. However, in *Pieris spp.*, evolution of NSP was a key innovation in the diversification of the genus as well as their ability to utilize Brassicaceae (Edger et al., 2015; Pomés et al., 1998; Wheat et al., 2007).

Apart from the above-mentioned specialist Lepidoptera, several species in the family Noctuidae (Lepidoptera) are known to feed on Brassicaceae in addition to feeding on a large range of hostplants. Several of these species are also well-known pest species including *Helicoverpa armigera*, *Spodoptera exigua* and *Trichoplusia ni* (Cho et al., 2008). Although they do not possess specialized enzymes mentioned above, they are able to feed on a diverse set of hostplant families including Brassicaceae through differential expression and regulation of the general detoxification machinery.

In addition to detoxification, sensory perception plays a crucial role in the hostplant range of insect herbivores. Sensory receptors are usually fine-tuned to the secondary metabolites present in the hostplant (Engsontia et al., 2014). For example, ovipositing females can differentiate between hosts and non-hosts as

well as assess the nutritional quality of the hostplant using gustatory and olfactory receptors (Ozaki et al., 2011; Ryuda et al., 2013). In fact, the diversity of sensory receptors is positively correlated with hostplant range in insects (Engsontia et al., 2014). Sensory receptors consist of three classes: The first class of receptors are olfactory receptors that aid in recognition of hostplants and conspecifics (Brand et al., 2018; Missbach et al., 2014). Gustatory receptors are involved in taste perception and consists of various subclasses that can identify sweet, bitter, and salt compounds (Nei et al., 2008; Sánchez-Gracia et al., 2009). The final class of sensory receptors are ionotropic receptors that are ligand gated channels that are crucial in the perception of secondary metabolites (Rytz et al., 2013).

Understanding the causal processes of hostplant specialization and the mechanisms through which generalist and specialist herbivores interact with well-defended plant toxins can shed light on the evolution and maintenance of detoxification machinery, which has important practical implications in agricultural management of pest species. Dynamic changes in gene family sizes, neofunctionalization and gene duplications in detoxification genes are hypothesized to be among the causal mechanism for the evolution of hostplant range in herbivorous insects (Dermauw et al., 2020; Dermauw & Van Leeuwen, 2014; Halon et al., 2015; Suzuki et al., 2018). Therefore, understanding the evolution of gene family size and selection on specialized detoxification genes can help shed light on the mechanisms through which polyphagous and specialist herbivores feed on well defended plant toxins.

We here use a comparative phylogenetic approach by using high quality genomes of lepidopteran insects that are both specialist and generalist feeders to test: a) if generalist herbivores have larger gene family changes in detoxification and sensory gene families compared to specialists; b) if detoxification and sensory gene families in generalist species evolve faster than specialists and c) if Nitrile Specifier Protein (NSP) was under purifying or diversifying selection.

Methods:

Data collection and quality assessment

We downloaded genomes of Lepidoptera which feed on Brassicaceae (table S1) from NCBI (Sayers et al., 2021) and Ensemble LepBase (Challi et al., 2016). The final list of species, the corresponding accession numbers and date accesses are provided in the supplementary table S1. We designated a species as a specialist herbivore if it only fed on Brassicaceae family and as a generalist herbivore if it fed on other plant families in addition to Brassicaceae. We assessed genome quality based on BUSCO scores with the Lepidoptera gene set consisting of 5218 BUSCO in BUSCO v. 4 (Simão et al., 2015). Genomes that showed a high percentage of fragmented or missing BUSCOs were excluded from downstream analysis.

Structural and functional annotation

We softmasked the repeats in our genomes using redmask (Girgis, 2015). Softmasked genomes were run through the BRAKER2 (Brůna et al., 2021) annotation pipeline with AUGUSTUS for gene predictions, with the Arthropoda

gene set from OrthoDB (Kriventseva et al., 2019) serving as the reference proteins. We used gffread from the cufflinks pipeline (Trapnell et al., 2012) to extract protein and CDS sequences following gene prediction by BRAKER2. We retained only one isoform per gene when genes were represented by multiple isoforms using agat_sp_keep_longest_isoform_pl function in AGAT (Dainat et al., 2020). The final protein set was cleaned of illegal characters such as "*" and "." and used for functional annotation with EggNOG (Cantalapiedra et al., 2021) and InterProScan (P. Jones et al., 2014). We used the Arthropoda dataset for functional annotation and realigned all the terms to their respective protein families using the Pfam (Mistry et al., 2021) databse for EggNOG annotation. For InterProScan, we used the -appl Pfam -goterms to get information about protein family annotations. We ran BUSCO again on the final protein set using the lepidopteran gene set to assess completeness of the proteome.

Orthologs inference and phylogenetic tree inference

We used OrthoFinder (Emms & Kelly, 2019) to predict orthologous protein groups (OGs). An orthogroup is a set of genes that have descended from a single gene from the last common ancestor (LCA) in a clade of species. The filtered protein set was used as input for Orthofinder and was run with default settings to predict OGs.

We additionally used Orthofinder for constructing our species and gene phylogenetic tree. We used the -M msa option in Orthofinder for multiple sequence alignment using MAFFT (Katoh & Standley, 2013). We used the

aligned single copy orthologs for building our species phylogenetic tree and the aligned sequences for each OG to construct the respective gene trees using IQTree (Nguyen et al., 2015). We used the default setting under IQTree for inferring species and gene trees. The root of the species tree was time calibrated to 87 Myr to reflect the divergence of the three families used in the analysis (Espeland et al., 2018).

Gene family evolution

CAFE v. 5.0 (Mendes et al., 2020) was used to analyze gene family evolution under a phylogenetic framework. We used the phylogenetic hierarchical orthogroups from OrthoFinder analysis to retrieve gene counts per species as input for CAFE. We first filtered OGs whose gene counts were greater than 100 in any single species as recommend by the developers. We additionally filtered OGs that showed high variance across species as they can lead to biases in gene family evolution estimation as well as preventing convergence among replicates in the analysis.

Once our test analysis using the base parameters showed convergence, we calculated the error in gene assembly and annotation in our dataset using the -e option in CAFE. We subsequently used the estimated error in our subsequent analysis to account for genome annotation errors. We used 4 different data sets for our gene family evolution analysis. Our first dataset included OGs comprised of all gene families after filtering of high variance and high copy numbers as indicated above, hereby referred to as "all genes". Second, we generated a

filtered dataset consisting of eight gene families involved in detoxification of plant compounds (p450 and COesterase in Step I, UGT, GST and ST in Step II, ABC transporter in Step III and insect cuticle protein and trypsin related to general digestive process), hereby referred to as "detoxification genes". Our third dataset included a filtered dataset consisting of three gene families involved in sensory recognition in Lepidoptera (gustatory, ionotropic, and olfactory genes), hereby referred to as "sensory genes". Our final filtered data set consisted of individual gene families involved in detoxification process and sensory recognition referred to as "single gene family".

We ran CAFE in two different modes. Our first run consisted of estimating a single rate of change (λ) for each of the 4 datasets to calculate a "baseline" λ for the entire tree. In our second run, we estimated λ separately for specialists, generalists and *Plutella xylostella*. We calculated a separate λ for *P. xylostella* for 2 reasons: a) *Plutella xylostella* was indicated as an outgroup in our tree, and therefore, including it with the rest of the specialists would bias our λ estimates for specialists and b) *P. xylostella* belongs to a different family compared to all the other specialist feeders (*Pieris sps.*, Pieridae) and generalist feeders (Noctuidae) in our dataset. We ran CAFE multiple times on all our datasets to ensure convergence among runs and used those estimates for determining λ for gene family evolution.

Signatures of selection on NSP gene

We used the HyPhy suite (Pond et al., 2005) on the webserver Datamonkey (Weaver et al., 2018) to test signatures of selection on the NSP gene. We first used the multiple aligned sequence of the OG corresponding to NSP from the OrthoFinder analysis and converted it to aligned CDS sequences using Pal2Nal (Suyama et al., 2006) to use as input in HyPhy. We used FEL (fixed effects likelihood) to investigate whether individual sites in the NSP gene were subjected to pervasive selection. FEL uses a ML method to infer nonsynonymous (dN) and synonymous (dS) substitution rate for each site. We ran FEL with 1000 parametric replicates and used an asymptotic chi-squared test to assess significance. To test if selection on the NSP gene in the *Pieris spp.* (specialist and butterfly) clade was different compared to the rest (generalists and/or moths) of the phylogeny, we used contrast-FEL with branches corresponding to *Pieris spp.* as foreground selection and the rest of the branches as background selection. We assessed significant differences in dN/dS based on a q-value threshold of <0.2. Finally, we used MEME (Mixed Effects Model of Evolution) to test if sites in the NSP gene were subjected to episodic selection.

Results:

Genome quality, gene family statistics and phylogeny

Our analysis of twelve Lepidoptera genomes included 4 specialist Brassicaceae herbivores from the butterfly family Pieridae (*Pieris brassicae*, *P. macdunnoughii*, *P. napi* and *P. rapae*), one specialist herbivore, *Plutella xylostella* from Plutellidae and seven generalist herbivores from the Noctuidae

(*Helicoverpa armigera*, *Helicoverpa zea*, *Mamestra configurata*, *Noctua pronuba*, *Trichoplusia ni*, *Spodoptera exigua* and *Phlogophora meticulosa*). BUSCO analysis of proteomes indicated that all the species had > 87.5% completeness with an average of 90.9% completeness, <5% average duplication, <1.5% average fragmentation and < 3.5% average missingness. The number of annotated protein coding genes ranged from 16733 in *P. napi* to 28532 proteins in *P. meticulosa* (fig. 4.1; mean: 20730 proteins). The number of annotated proteins was higher in moths (families Noctuidae and Plutellidae; median: 20557) compared to butterflies (family Pieridae: 17053 proteins). The number of functionally annotated proteins ranged from 12840 to 17850 (mean: 14650 annotations).

OrthoFinder analysis identified 21785 orthogroups. The gene counts corresponding to these orthogroups were used as input for CAFE analysis. Our filtered OG count after removing high variance OGs among species and >100 genes in any single species included 20897 OGs in the "all genes" dataset. The "detoxification genes" dataset consisted of 690 OGs and the "sensory genes" dataset consisted of 69 OGs. The "single gene family" dataset consisted of the following number of OGs for each family: ABCs: 57, COesterases: 77, Insect cuticle proteins: 123, GSTs: 33, Gustatory genes: 6, Ionotropic genes: 14, Olfactory genes: 49, p450s: 101, Sulfotransferases: 28, Trypsins: 237 and finally, UGTs: 34.

Our species level phylogenetic tree was constructed based on 3575 complete single copy orthologs. Our phylogeny placed *Plutella xylostella* as the

outgroup and Pieridae and Noctuidae as sister groups which is consistent with other published phylogenies (fig. 4.1).

Gene family expansions and contractions

Our analysis suggested that across the entire phylogeny, *Mamestra configurata* had the highest gene family expansions for the "all genes" dataset (n=1873) and *Trichoplusia ni* had the highest gene family contractions (n=1269) (fig. 4.2). Among butterflies, *Pieris napi* had the highest gene family expansions (n=894) and gene family contractions (n=995) for the "all genes" dataset. Among the detoxification and sensory gene families, *M. configurata* had the highest gene family expansions for both families and the node leading to the *Pieris sps.* clade had the largest gene family contraction for detoxification genes and *P. napi* for the sensory gene families (fig. 4.2).

Gene family evolution:

Our analysis of gene family evolution for the "all genes" dataset resulted in an overall change, λ, of 0.00274 (-logL: 114663; fig. 4.3). Separate λ estimations for "all genes" showed that gene family evolution in *P. xylostella* (λ= 0.00058) was lower than the baseline change, while both specialists (λ=0.0027) and generalists (λ=0.0036) had a higher rate of evolution than baseline, with generalists having the highest rate of gene evolution. For "detoxification genes", generalist herbivores had higher rate of gene evolution (λ= 0.0071) compared to specialist herbivores (λ= 0.0062) and for "sensory genes", specialist herbivores had higher gene evolution rate (λ= 0.0059) than generalist herbivores.

In our "single gene family" dataset, *P. xylostella* had the lowest rate of evolution across all detoxification enzymes. In genes that constitute the first step in detoxification of plant compounds, generalist herbivores had the highest rate of change in p450 genes (λ=0.011) and in COesterase gene family (λ=0.0087). In the genes that facilitate conversion of plant toxic compounds to hydrophilic compounds, specialist herbivores had the highest rate of evolution in UGT (λ=0.0123), GST (λ=0.0087) and ST (λ=0.004). Generalist herbivores had higher rate of evolution for ABC (λ=0.0028) (which is involved in excreting toxic compounds), for trypsin (λ=0.007) and cuticle protein (λ=0.0058) (both involved in increasing digestive efficiency).

Our analysis of "single gene family" olfactory gene families indicated that gustatory gene evolution was highest in *P. xylostella* (λ=0.0043) and lowest in specialist herbivores (λ=0.0013) (*Pieris spp.*) while ionotropic and olfactory gene families evolved at a faster rate in specialist herbivores compared generalist feeders or *P. xylostella* (fig. 4.3).

Signatures of selection on NSP gene

We found evidence for pervasive purifying selection in the NSP gene (fig.4.4). 41% of the codons were under purifying selection (95/227 codons). Our analysis of dN/dS showed that the NSP gene was characterized predominantly by synonymous substitutions. Almost all of the codons under purifying selection were present in the conserved domain of the NSP gene. However, we did not identify any sites that were under pervasive diversifying/positive selection. We

used Contrast-FEL to test if the NSP gene in the *Pieris* genus had different substitution rates compared to the rest of the phylogeny. We identified 2 sites (blue arrow in fig. 4.4) that showed different substitution rates in the *Pieris spp.* clade compared to the rest of the phylogeny. Additionally, we identified six sites in the conserved domain that were under episodic positive/diversifying selection (black arrow in fig. 4.4).

Discussion:

We evaluated the role of detoxification and sensory gene family expansions in specialist and generalist herbivores feeding on Brassicaceae. We additionally examined selection on the nitrile specifier protein gene, a novel protein involved in glucosinolate detoxification in Lepidoptera. We found that generalist herbivores had the largest gene family changes in all gene families, detoxification related families as well as sensory gene families. Generalist herbivores also had faster evolution of all genes and detoxification gene families, while specialists had faster gene evolution in sensory gene families. Additionally, we found that the nitrile specifier protein gene is highly conserved among generalist and specialist lepidopterans. Although NSP is highly conserved, we identified signatures of episodic and positive selection on the gene.

Gene family expansion/contraction in Lepidoptera

We examined gene family expansions/contractions in 12 species of Lepidoptera from 3 distinct families that feed on Brassicaceae. Although many other Lepidoptera and insects are known to feed on Brassicaceae, the 12

species included in our analysis are known to cause serious damage to Brassicaceae. In fact, species in the Noctuidae family constitute some of the most damaging pests in the world (Cho et al., 2008). Genera such as *Spodoptera*, *Trichoplusia* and *Helicoverpa* are some of the insects with the widest range of recorded hostplants (G. S. Robinson et al., 2010). All species in the Noctuidae family in our analysis are polyphagous and also feed on Brassicaceae, while Pieridae and Plutellidae family in our analysis consists of specialist herbivores on Brassicaceae.

We found that Noctuidae family and the species represented within showed the largest gene family expansions and contractions, followed by *Plutella xylostella* (fig. 4.2). *Pieris sps.* had the lowest gene family expansion. Our results, therefore, suggest that polyphagy is associated with greater gene family expansions and contractions. This is in line with other studies that have found that species with larger hostplant ranges have large gene families compared to those that fed on few or a single plant family (Breeschoten, van der Linden, et al., 2022; Suzuki et al., 2018).

Polyphagous herbivores use the general detoxification processes (phase I -III) to detoxify plant compounds and thus, larger copy numbers of these genes will result in higher gene expression to process plant toxins efficiently (Breeschoten, van der Linden, et al., 2022; Dermauw & Van Leeuwen, 2014; Ranson et al., 2002). Polyphagy is also associated with a dynamic gene family expansion/contraction, especially if the hostplant range of the herbivore is labile. For example, in the Nymphalidae, hostplant range oscillates from ancestral

specialists that experience range extension giving rise to generalist feeders that adapt to local hostplant communities and in turn diversify and specialize (Hardy & Otto, 2014; Janz et al., 2006; Janz & Nylin, 2008; Nylin et al., 2014). This results in rapid changes in gene family expansions/contractions. If hostplant specialization is selected in the future, this can result in fewer gene family expansion/contractions. In fact, that is what we see in specialist Brassicaceae feeders. *Pieris spp.* and *P. xylostella* on the other hand are specialist herbivores that possess fine-tuned detoxification processes to counteract glucosinolate toxicity. Therefore, generalized detoxification genes families are relatively stable and under strong selection.

Gene evolution in Lepidoptera

Our CAFE analysis estimated an overall change of 0.00274 (gains/losses/Myr) for all genes. Our estimate of λ is consistent with what is seen in other insects such as *Drosophila*, λ= 0.0012; (M. W. Hahn et al., 2007), Lepidoptera λ= 0.0023; (Breeschoten, van der Linden, et al., 2022), and *Anopheles* λ= 0.0031 (Neafsey et al., 2015). When estimating separate change for specialists, generalists, and *P. xylostella*, we found a higher rate of change for generalists, which fits with the largest gene family expansion/contraction in generalist species (Breeschoten, van der Linden, et al., 2022; Suzuki et al., 2018). Higher gene family evolution rates could be due to underlying gene duplication, neo-functionalization and/or genome rearrangements, all of which are implicated in polyphagous feeding (Murad et al., 2021; Seppey et al., 2019).

Our estimation of gene evolution for detoxification gene families showed that generalists had the highest rate of change. When we examined individual detoxification gene families, generalists had higher gene evolution for phase I, phase III and digestion related genes (cuticle and trypsin). p450 genes have been extensively studied for their role in hostplant use in Lepidoptera. p450 genes can metabolize a wide range of plant compounds including xanthoxin (Mao et al., 2006), furanocumarins (Lindroth, 1989), alkaloids (Wang et al., 2018) and even insecticide resistance (Pan et al., 2018). Since polyphagous herbivores feed on a wide range of chemicals, faster evolution in genes of phase I detoxification can help them adapt to new hosts and overcome insecticide resistance. Gene families in phase II had a higher evolutionary rate in specialist herbivores compared to generalists. Phase II detoxification genes are specialized genes that add specific compounds which increase solubility and excretion efficiency (Nallu et al., 2018; You et al., 2015). Specialist herbivores have alternate pathways to detoxify toxic compounds, for example, the use of NSP to form nitriles instead of isothiocyanates in specialist Brassicaceae feeders (Edger et al., 2015; Okamura, Sato, Tsuzuki, Sawada, et al., 2019; Ratzka et al., 2002). These modified products are processed by phase II genes and specialist herbivores have evolved modified phase II enzymes to efficiently process these by products. In fact, GSTs are upregulated in the larvae of *Pieris spp.* and *Plutella xylostella* when feeding on Brassicaceae (You et al., 2015). UDPGTs are also enriched in the transcriptome of *Danus plexippus* larvae which feed on plants from the family Apocynaceae which contain cardiac glycosides (Ranz et

al., 2021). Generalist herbivores had higher evolution in phase III as well trypsin and insect cuticle proteins, all of which are positively correlated with polyphagy in many insect species including Lepidoptera (Kelkenberg et al., 2015; Muhlia-Almazán et al., 2008; Rawlings & Barrett, 1995).

Our estimation of gene evolution in sensory receptors highlighted that overall, specialist herbivores had faster evolution compared to generalists. Our single gene family analysis, however showed faster evolution of gustatory receptors in *P. xylostella* compared to generalists or specialists. This is in agreement with what has been published for lepidopteran gustatory receptors. *Plutella xylostella* is an outlier and possess greater diversity of gustatory receptors and is known to have higher gene evolution compared to other lepidopteran species (Engsontia et al., 2014). The gene evolution of gustatory gene families is positively correlated with host plant range. For example, in a transcriptomic analysis of gustatory gene evolution in Nymphalidae, the lineage leading to *Vanessa cardui*, a generalist herbivore known to feed on over 400 hostplants had the highest gene evolution rate compared to other specialists (Suzuki et al., 2018). Our analysis of ionotropic and olfactory gene evolution suggested higher rate of evolution in specialist herbivores compared to generalist herbivores. Specialist herbivores feeding on Brassicaceae have fine-tuned olfactory receptors to identify glucosinolates that have evolved through gene duplication and neofunctionalization and therefore expected to have higher evolution compared to those in generalist herbivores (Matsunaga et al., 2022).

Selection on NSP gene

Our analysis of selection on the NSP gene in Brassicaceae-feeding Lepidopteran insects revealed strong purifying selection. Our dN/dS estimates indicated a lack of non-synonymous substitutions, suggesting the absence of pervasive diversifying selection on NSP. Our contrast-FEL analysis showed that the lineage leading to *Pieris spp.* had 2 sites in the conserved domain that had different substitution rates compared to the other species. This could indicate selection specific to *Pieris spp.* on the NSP gene. Our results are in agreement with what has been shown for positive selection on NSP gene in a subgroup of *Pieris spp*. (Okamura, Sato, Tsuzuki, Murakami, et al., 2019). The neofunctionalization and cooption of the insect allergen protein in *Pieris* spp. to the extant NSP gene is a key innovation that allowed *Pieris spp.* to colonize Brassicaceae and diversify on them. The two codons that we identified could represent the differential selection pressures experienced by *Pieris spp* compared to the rest of the phylogeny. Finally, our test for episodic selection identified 6 sites that were subjected to diversifying selection. Although our analysis does not shed light on the functional role of the sites under selection, it does highlight the likely genetic targets that are crucial for adaptation and feeding on Brassicaceae. These sites could serve as the targets for manipulative CRISPR experiments to understand the mechanisms through which specialist herbivores detoxify glucosinolates in Brassicaceae.

Genome quality and Lepidopteran phylogenetic framework

Our interpretations of gene family evolution and selection on the NSP gene is critically dependent on the accuracy of genome annotations as well as

species phylogeny. The availability of genomes for non-model organisms has increased considerably with the advent of affordable sequencing and development of assembly and annotation pipelines. However, the quality of assembly and annotation varies considerably due to differences in sequencing methods as well as differences in analytical methods among various research groups (Ellis et al., 2021). To control for variation among different genomes, we employed several strict quality checks and additional analysis: 1) We checked the quality of the genome using BUSCO and used only those genomes that had a BUSCO complete, single copy score >80%. 2.) We used BRAKER 2 for gene prediction and EggNOG for functional annotation instead of using the default proteome files available with the genomes. The annotated proteomes were assessed for quality using BUSCO and only those that had a BUSCO complete single copy score >85% were retained. This ensured that all proteomes were of comparable quality for downstream analysis. 3) We retained only the longest isoform per gene and removed isoform duplications that could bias gene family evolution estimates. 4) We assigned a separate lambda for Plutella xylostella since it was the outgroup in our analysis as well as a specialist Lepidoptera belonging to a different family from the other species. 5) We calculated the error in CAFE analysis and applied that to our analysis to account for genome annotation and assembly errors.

Our phylogenetic reconstruction using single copy orthologs placed *P. xylostella* as the outgroup with Noctuidae and Pieridae as sister genera which is consistent with published reports (Espeland et al., 2018). We used 87 MYA as

the root calibration, the age of diversification of *P. xylostella*. This placed the age of node split of *P. napi* and *P. macdunnoughii* about 4-6 MYA, which is consistent with what the Holarctic expansion and diversification of *P. napi* (Geiger & Shapiro, 1992).

Conclusion:

We used whole genome comparative analysis to examine the role of detoxification gene families, sensory gene families and the specialized NSP gene on Lepidopteran herbivory on Brassicaceae. We found positive correlations with general detoxification gene expansions and polyphagy and identified specific steps in the detoxification pathway where gene family expansions were correlated with host specialization. We additionally identified signatures of pervasive purifying selection on the NSP gene as well as identified sites that were subjected to different selection pressures in the lineage leading to *Pieris* spp. Our results add to the growing body of work on understanding gene family evolution and its role in hostplant range and specialization in insects.

Acknowledgements:

We thank Ward B Watt, Joseph M Quattro, Brian Hollis, and Peter Andolfatto for proving comments on earlier versions of the manuscript. We thank Muktai Kuwalekar, and Hannah Dort for advice on gene family evolution analysis. This work was supported by the University of South Carolina Arts & Sciences to CLB.

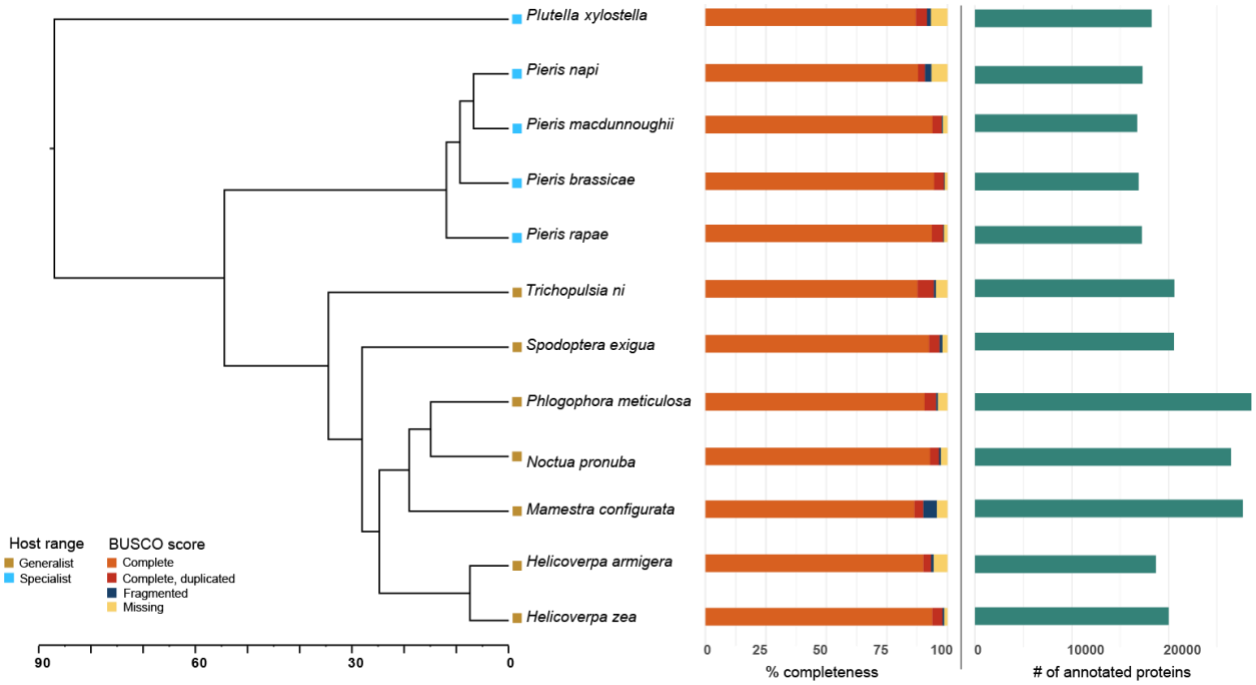

Figure 4.1: ML tree based on single copy orthologs from 12 Lepidopteran species. Host range legend reflects herbivore specialist/generalist status. BUSCO score represents protein completeness.

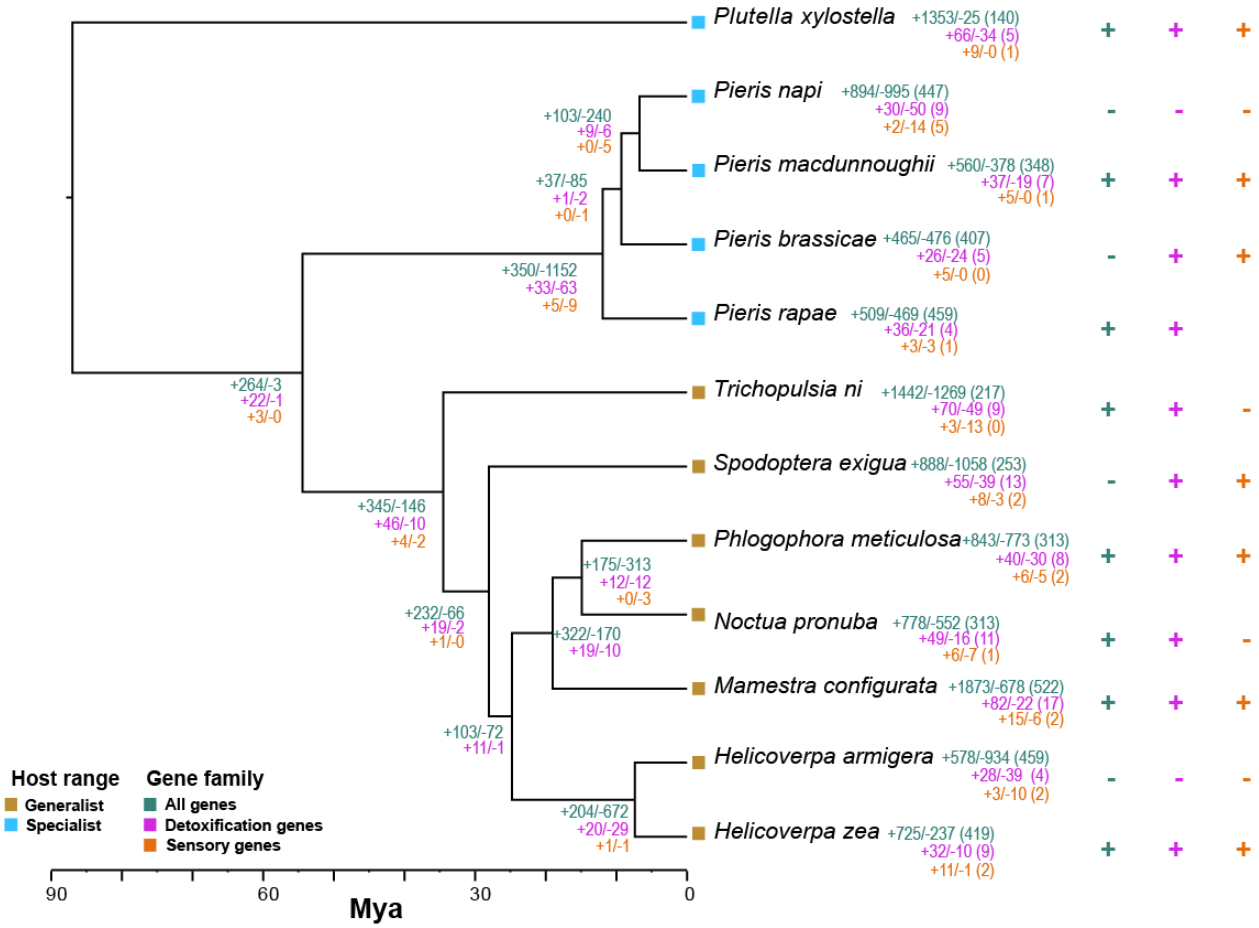

Figure 4.2: ML tree showing gene family expansion (indicated as +) and contraction (indicated as -) for 12 species. Gene family color legend reflects expansion/contractions for the corresponding gene family.

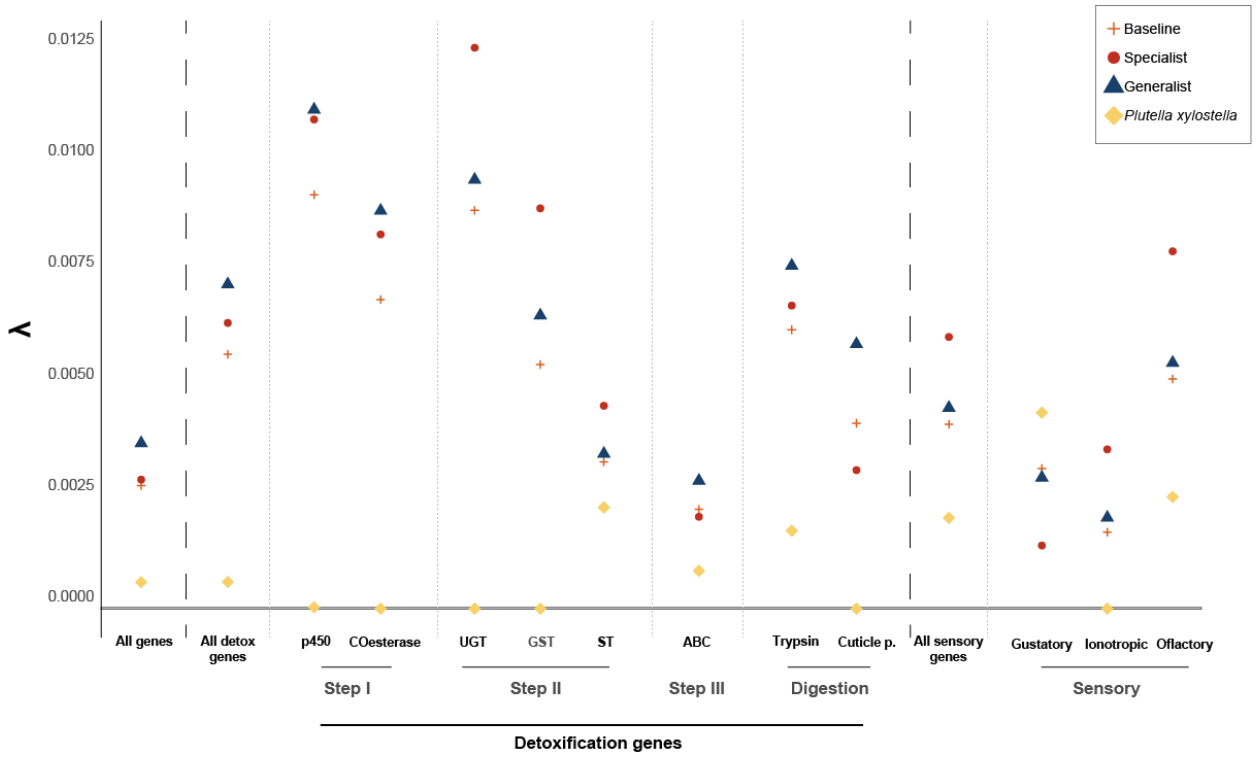

Figure 4.3: Gene family evolution rates calculated from CAFE. Rates reflect change/gene/Myr.

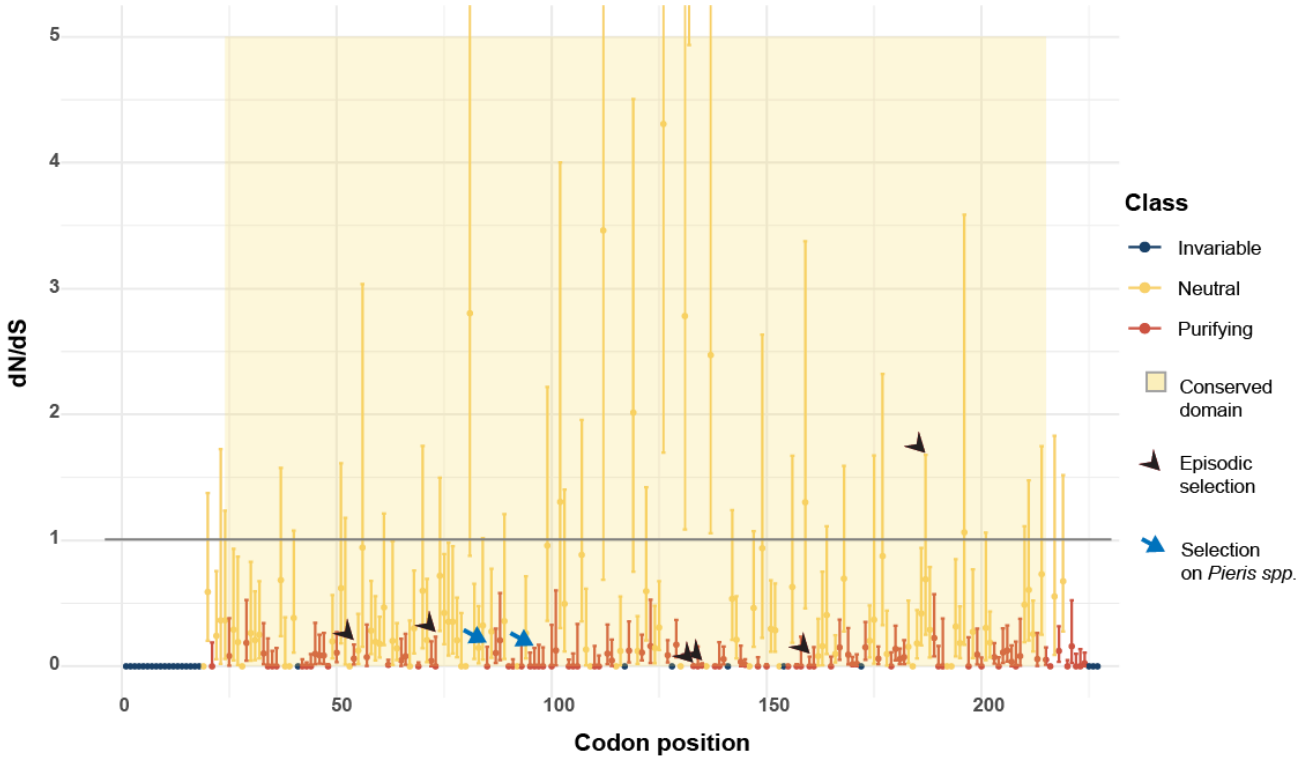

Figure 4.4: Molecular evolution of NSP gene. Episodic selection and differential selection highlighted in black and blue arrows respectively.

REFERENCES

Adiba, S., Huet, M., & Kaltz, O. (2010). Experimental evolution of local parasite maladaptation. *Journal of Evolutionary Biology*, *23*(6), 1195–1205. https://doi.org/10.1111/j.1420-9101.2010.01985.x

Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Al-Asadi, H., Petkova, D., Stephens, M., & Novembre, J. (2018). Estimating recent migration and population size surfaces. *BioRxiv*, 1–21. https://doi.org/10.1101/365536

Alexa, A., & Rahnenführer, J. (2009). Gene set enrichment analysis with topGO. *Bioconductor Improv*, *27*, 1–26.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Alleaume-Benharira, M., Pen, I. R., & Ronce, O. (2006). Geographical patterns of adaptation within a species' range: Interactions between drift and gene flow. *Journal of Evolutionary Biology*, *19*(1), 203–215. https://doi.org/10.1111/j.1420-9101.2005.00976.x

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson, K. V. (2000). Toll signaling pathways in the innate immune response. *Current Opinion in Immunology*, *12*(1), 13–19. https://doi.org/10.1016/S0952-7915(99)00045-X

Andolfatto, P. (2005). Adaptive evolution of non-coding DNA in *Drosophila*. *Nature*, *437*(7062), 1149.

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T. T., Mast, J., Sunayama-Morita, T., & Stern, D. L. (2011). Multiplexed shotgun genotyping for rapid and

efficient genetic mapping. *Genome Research*, *21*(4), 610–617. https://doi.org/10.1101/gr.115402.110

Andrews, S. (2010). *A quality control tool for high throughput sequence data.* http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Austen, E. J., Rowe, L., Stinchcombe, J. R., & Forrest, J. R. K. (2017). Explaining the apparent paradox of persistent selection for early flowering. *New Phytologist*, *215*(3), 929–934. https://doi.org/10.1111/nph.14580

Awmack, C. S., & Leather, S. R. (2002). Host Plant Quality and Fecundity in Herbivorous Insects. *Annual Review of Entomology*, *47*(1), 817–844. https://doi.org/10.1146/annurev.ento.47.091201.145300

Bachmann, J. C., Rensburg, A. J. van, Cortazar-Chinarro, M., Laurila, A., & Buskirk, J. V. (2020). Gene flow limits adaptation along steep environmental gradients. *American Naturalist*, *195*(3), E67–E86. https://doi.org/10.1086/707209

Barton, N., & Partridge, L. (2000). Limits to natural selection. *BioEssays*, *22*(12), 1075–1084. https://doi.org/10.1002/1521-1878(200012)22:12<1075::AID-BIES5>3.0.CO;2-M

Beerli, P., & 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*(8), 4563–4568. https://doi.org/10.1073/pnas.081068098

Bell, G. (2017). Evolutionary Rescue. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 605–627. https://doi.org/10.1146/annurev-ecolsys-110316-023011

Bell, G., & Gonzalez, A. (2011). Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science*, *332*(6035), 1327–1330. https://doi.org/10.1126/science.1203105

Berdegué, M., Reitz, S. R., & Trumble, J. T. (1998). Host plant selection and development in *Spodoptera exigua*: Do mother and offspring know best? *Entomologia Experimentalis et Applicata*, *89*(1), 57–64. https://doi.org/10.1046/j.1570-7458.1998.00381.x

Berenbaum, M. (1983). Coumarins and Caterpillars: A Case for Coevolution. *Evolution*, *37*(1), 163–179. https://doi.org/10.2307/2408184

Berenbaum, M. R., Zangerl, A. R., & Nitao, J. K. (1986). Constraints on Chemical Coevolution: Wild Parsnips and the Parsnip Webworm. *Evolution*, *40*(6), 1215–1228. https://doi.org/10.1111/j.1558-5646.1986.tb05746.x

Bernays, E., & Graham, M. (1988). On the Evolution of Host Specificity in Phytophagous Arthropods. *Ecology*, *69*(4), 886–892. https://doi.org/10.2307/1941237

Best, K. F., & McIntyre, G. I. (1975). The biology of Canadian weeds: 9. *Thlaspi arvense* L. *Canadian Journal of Plant Science*, *55*(1), 279–292. https://doi.org/10.4141/cjps75-039

Biddanda, A., Rice, D. P., & Novembre, J. (2020). A variant-centric perspective on geographic patterns of human allele frequency variation. *ELife*, *9*, e60107. https://doi.org/10.7554/eLife.60107

Bird, C. P., Stranger, B. E., & Dermitzakis, E. T. (2006). Functional variation and evolution of non-coding DNA. *Current Opinion in Genetics & Development*, *16*(6), 559–564. https://doi.org/10.1016/j.gde.2006.10.003

Black, T. V., & Robertson, B. A. (2020). How to disguise evolutionary traps created by solar panels. *Journal of Insect Conservation*, *24*(2), 241–247. https://doi.org/10.1007/s10841-019-00191-5

Blankenchip, C. L., Michels, D. E., Braker, H. E., & Goffredi, S. K. (2018). Diet breadth and exploitation of exotic plants shift the core microbiome of *Cephaloleia*, a group of tropical herbivorous beetles. *PeerJ*, *6*, e4793. https://doi.org/10.7717/peerj.4793

Blanquart, F., Gandon, S., & Nuismer, S. L. (2012). The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, *25*(7), 1351–1363. https://doi.org/10.1111/j.1420-9101.2012.02524.x

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013a). A practical guide to measuring local adaptation. In *Ecology Letters* (Vol. 16, Issue 9, pp. 1195–1205). https://doi.org/10.1111/ele.12150

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013b). A practical guide to measuring local adaptation. *Ecology Letters*, *16*(9), 1195–1205. https://doi.org/10.1111/ele.12150

Boag, P. T., & Grant, P. R. (1981). Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galápagos. *Science*, *214*(4516), 82–85. https://doi.org/10.1126/science.214.4516.82

Boggs, C. L. (1981). Nutritional and life-history determinants of resource allocation in holometabolous insects. *The American Naturalist*, *117*(5), 692–709. https://doi.org/10.2307/2678832

Boggs, C. L. (2009). Understanding insect life histories and senescence through a resource allocation lens. *Functional Ecology*, *23*(1), 27–37. https://doi.org/10.1111/j.1365-2435.2009.01527.x

Boggs, C. L., & Freeman, K. D. (2005). Larval food limitation in butterflies: Effects on adult resource allocation and fitness. *Oecologia*, *144*(3), 353–361. https://doi.org/10.1007/s00442-005-0076-6

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bolnick, D. I., Caldera, E. J., & Matthews, B. (2008). Evidence for asymmetric migration load in a pair of ecologically divergent stickleback populations. *Biological Journal of the Linnean Society*, *94*(2), 273–287. https://doi.org/10.1111/j.1095-8312.2008.00978.x

Bolnick, D. I., & Nosil, P. (2007). Natural selection in populations subject to a migration load. *Evolution*, *61*(9), 2229–2243. https://doi.org/10.1111/j.1558-5646.2007.00179.x

Bossart, J. L., & Scriber, J. M. (1995). Maintenance of ecologically significant genetic variation in the tiger swallowtail butterfly through differential selection and gene flow. *Evolution*, *49*(6), 1163–1171. https://doi.org/10.1111/j.1558-5646.1995.tb04443.x

Bossart, J. L., & Scriber, J. M. (1999). Preference Variation in the Polyphagous Tiger Swallowtail Butterfly (Lepidoptera: Papilionidae). *Environmental Entomology*, *28*(4), 628–637. https://doi.org/10.1093/ee/28.4.628

Bradshaw, A. D., & McNeilly, T. (1991). Evolutionary response to global climatic change. *Annals of Botany*, *67*(SUPPL. 1), 5–14. https://doi.org/10.1093/oxfordjournals.aob.a088209

Brady, S. P., Bolnick, D. I., Angert, A. L., Gonzalez, A., Barrett, R. D. H., Crispo, E., Derry, A. M., Eckert, C. G., Fraser, D. J., Fussmann, G. F., Guichard, F., Lamy, T., McAdam, A. G., Newman, A. E. M., Paccard, A., Rolshausen, G., Simons, A. M., & Hendry, A. P. (2019). Causes of maladaptation. *Evolutionary Applications*, *12*(7), 1229–1242. https://doi.org/10.1111/eva.12844

Brady, S. P., Bolnick, D. I., Barrett, R. D. H., Chapman, L., Crispo, E., Derry, A. M., Eckert, C. G., Fraser, D. J., Fussmann, G. F., Gonzalez, A., Guichard, F., Lamy, T., Lane, J., McAdam, A. G., Newman, A. E. M., Paccard, A., Robertson, B., Rolshausen, G., Schulte, P. M., … Hendry, A. P. (2019). Understanding maladaptation by uniting ecological and evolutionary perspectives. *American Naturalist*, *194*(4), 495–515. https://doi.org/10.1086/705020

Brand, P., Robertson, H. M., Lin, W., Pothula, R., Klingeman, W. E., Jurat-Fuentes, J. L., & Johnson, B. R. (2018). The origin of the odorant receptor gene family in insects. *ELife*, *7*, e38340. https://doi.org/10.7554/eLife.38340

Breeschoten, T., Ros, V. I. D., Schranz, M. E., & Simon, S. (2019). An influential meal: Host plant dependent transcriptional variation in the beet armyworm, *Spodoptera exigua* (Lepidoptera: Noctuidae). *BMC Genomics*, *20*(1), 845. https://doi.org/10.1186/s12864-019-6081-7

Breeschoten, T., Schranz, M. E., Poelman, E. H., & Simon, S. (2022). Family dinner: Transcriptional plasticity of five Noctuidae (Lepidoptera) feeding on three host plant species. *Ecology and Evolution*, *12*(9), e9258. https://doi.org/10.1002/ece3.9258

Breeschoten, T., van der Linden, C. F. H., Ros, V. I. D., Schranz, M. E., & Simon, S. (2022). Expanding the Menu: Are Polyphagy and Gene Family Expansions Linked across Lepidoptera? *Genome Biology and Evolution*, *14*(1). https://doi.org/10.1093/gbe/evab283

Bridle, J. R., Kawata, M., & Butlin, R. K. (2019). Local adaptation stops where ecological gradients steepen or are interrupted. *Evolutionary Applications*, *12*(7), 1449–1462. https://doi.org/10.1111/eva.12789

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2021). BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics*, *3*(1), lqaa108. https://doi.org/10.1093/nargab/lqaa108

Calla, B., Noble, K., Johnson, R. M., Walden, K. K. O., Schuler, M. A., Robertson, H. M., & Berenbaum, M. R. (2017). Cytochrome P450 diversification and hostplant utilization patterns in specialist and generalist moths: Birth, death and adaptation. *Molecular Ecology*, *26*(21), 6021–6035. https://doi.org/10.1111/mec.14348

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829. https://doi.org/10.1093/molbev/msab293

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-alonso, M., & Keller, S. R. (2020). Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*, 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Caprio, M. A., & Tabashnik, B. E. (1992). Gene Flow Accelerates Local Adaptation Among Finite Populations: Simulating the Evolution of Insecticide Resistance. *Journal of Economic Entomology*, *85*(3), 611–620. https://doi.org/10.1093/jee/85.3.611

Carlson, S. M., Cunningham, C. J., & Westley, P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology and Evolution*, *29*(9), 521–530. https://doi.org/10.1016/j.tree.2014.06.005

Carlson, S. M., & Quinn, T. P. (2007). Ten years of varying lake level and selection on size-at-maturity in sockeye salmon. *Ecology*, *88*(10), 2620–2629. https://doi.org/10.1890/06-1171.1

Carrasco, D., Larsson, M. C., & Anderson, P. (2015). Insect host plant selection in complex environments. *Current Opinion in Insect Science*, *8*, 1–7. https://doi.org/10.1016/j.cois.2015.01.014

Casagrande, R. A., & Dacey, J. E. (2007). Monarch Butterfly Oviposition on Swallow-Worts (*Vincetoxicum spp.*). *Environmental Entomology*, *36*(3), 631–636. https://doi.org/10.1603/0046-225X(2007)36[631:MBOOSV]2.0.CO;2

Castledine, M., Padfield, D., & Buckling, A. (2020). Experimental (co)evolution in a multi-species microbial community results in local maladaptation. *Ecology Letters*, *23*(11), 1673–1681. https://doi.org/10.1111/ele.13599

Ceballos, G., & Ehrlich, P. R. (2002). Mammal population losses and the extinction crisis. *Science*, *296*(5569), 904–907. https://doi.org/10.1126/science.1069349

Celorio-Mancera, M. de la P., Wheat, C. W., Huss, M., Vezzi, F., Neethiraj, R., Reimegård, J., Nylin, S., & Janz, N. (2016). Evolutionary history of host use, rather than plant phylogeny, determines gene expression in a generalist butterfly. *BMC Evolutionary Biology*, *16*(1), 59. https://doi.org/10.1186/s12862-016-0627-y

Cermak, R. (2008). Effect of dietary flavonoids on pathways involved in drug metabolism. *Expert Opinion on Drug Metabolism & Toxicology*, *4*(1), 17–35. https://doi.org/10.1517/17425255.4.1.17

Challi, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D., & Blaxter, M. (2016). Lepbase: The Lepidopteran genome database. *BioRxiv*. https://doi.org/10.1101/056994

Charlesworth, D. (2006). Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, *2*(4), 379–384. https://doi.org/10.1371/journal.pgen.0020064

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chew, F. S. (1977a). Coevolution of Pierid Butterflies and Their Cruciferous Foodplants. II. The Distribution of Eggs on Potential Foodplants. *Evolution*, *31*(3), 568–579. https://doi.org/10.2307/2407522

Chew, F. S. (1977b). The Effects of Introduced Mustards (Cruciferae) on Some Native North American Cabbage Butterflies (Lepidoptera: Pieridae). *Atala*, *5*(2), 13–19.

Chew, F. S. (1980). Foodplant preferences of Pieris caterpillars (Lepidoptera). *Oecologia*, *46*(3), 347–353. https://doi.org/10.1007/BF00346263

Chew, F. S., & Watt, W. B. (2006). The green-veined white (*Pieris napi* L.), its Pierine relatives, and the systematics dilemmas of divergent character sets (Lepidoptera, Pieridae). *Biological Journal of the Linnean Society*, *88*(3), 413–435. https://doi.org/10.1111/j.1095-8312.2006.00630.x

Cho, S., Mitchell, A., Mitter, C., Regier, J., Matthews, M., & Robertson, R. (2008). Molecular phylogenetics of heliothine moths (Lepidoptera: Noctuidae: Heliothinae), with comments on the evolution of host range and pest status. *Systematic Entomology*, *33*(4), 581–594. https://doi.org/10.1111/j.1365-3113.2008.00427.x

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Clissold, F. J., & Simpson, S. J. (2015). Temperature, food quality and life history traits of herbivorous insects. *Current Opinion in Insect Science*, *11*, 63–70. https://doi.org/10.1016/j.cois.2015.10.011

Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, *24*(9), 1999–2017. https://doi.org/10.1111/mec.13162

Cotto, O., Sandell, L., Chevin, L. M., & Ronce, O. (2019). Maladaptive shifts in life history in a changing environment. *American Naturalist*, *194*(4), 558–573. https://doi.org/10.1086/702716

Crespi, B. J. (2000). The evolution of maladaptation. *Heredity*, *84*(6), 623–629. https://doi.org/10.1046/j.1365-2540.2000.00746.x

Crow, J. F. (1970). Genetic Loads and the Cost of Natural Selection. In *Mathematical Topics in Population Genetics* (pp. 128–177). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-46244-3_5

Dainat, J., Hereñú, D., & Pucholt, P. (2020). AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF. *GFF Format. Zenodo*.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray.

de Fouchier, A., Walker, W. B., Montagné, N., Steiner, C., Binyameen, M., Schlyter, F., Chertemps, T., Maria, A., François, M.-C., Monsempes, C., Anderson, P., Hansson, B. S., Larsson, M. C., & Jacquin-Joly, E. (2017). Functional evolution of Lepidoptera olfactory receptors revealed by deorphanization of a moth repertoire. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms15709

Delibes, M., Gaona, P., & Ferreras, P. (2001). Effects of an attractive sink leading into maladaptive habitat selection. *American Naturalist*, *158*(3), 277–285. https://doi.org/10.1086/321319

Dermauw, W., & Van Leeuwen, T. (2014). The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. *Insect Biochemistry and Molecular Biology*, *45*, 89–110. https://doi.org/10.1016/j.ibmb.2013.11.001

Dermauw, W., Van Leeuwen, T., & Feyereisen, R. (2020). Diversity and evolution of the P450 family in arthropods. *Insect Biochemistry and Molecular Biology*, *127*, 103490. https://doi.org/10.1016/j.ibmb.2020.103490

Derry, A. M., Fraser, D. J., Brady, S. P., Astorg, L., Lawrence, E. R., Martin, G. K., Matte, J. M., Negrín Dastis, J. O., Paccard, A., Barrett, R. D. H., Chapman, L. J., Lane, J. E., Ballas, C. G., Close, M., & Crispo, E. (2019). Conservation through the lens of (mal)adaptation: Concepts and meta-analysis. *Evolutionary Applications*, *12*(7), 1287–1304. https://doi.org/10.1111/eva.12791

Diamond, S. E., & Kingsolver, J. G. (2010). Host plant quality, selection history and trade-offs shape the immune responses of *Manduca sexta*. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1703), 289–297. https://doi.org/10.1098/rspb.2010.1137

Diamond, S. E., & Martin, R. A. (2020). Evolution is a double-edged sword, not a silver bullet, to confront global change. *Annals of the New York Academy of Sciences*, *1469*(1), 38–51. https://doi.org/10.1111/nyas.14410

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dobin, A., & Gingeras, T. R. (2015). Mapping RNA-seq Reads with STAR. *Current Protocols in Bioinformatics*, *51*, 11.14.1--11.14.19. https://doi.org/10.1002/0471250953.bi1114s51

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glöckner, G., Platts, A. E., Heckel, D. G., Der, J. P., Wafula, E. K., Tang, M., Hofberger, J. A., Smithson, A., Hall, J. C., Blanchette, M., Bureau, T. E., Wright, S. I., dePamphilis, C. W., Eric Schranz, M., Barker, M. S., … Wheat, C. W. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences*, *112*(27), 8362–8366. https://doi.org/10.1073/pnas.1503926112

Egan, S. P., Ragland, G. J., Assour, L., Powell, T. H. Q., Hood, G. R., Emrich, S., Nosil, P., & Feder, J. L. (2015). Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. *Ecology Letters*, *18*(8), 817–825. https://doi.org/10.1111/ele.12460

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and Plants: A Study in Coevolution. *Evolution*, *18*(4), 586–608.

Ellis, E. A., Storer, C. G., & Kawahara, A. Y. (2021). De novo genome assemblies of butterflies. *GigaScience*, *10*(6), 1–8. https://doi.org/10.1093/gigascience/giab041

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Engsontia, P., Sangket, U., Chotigeat, W., & Satasook, C. (2014). Molecular evolution of the odorant and gustatory receptor genes in lepidopteran insects: Implications for their adaptation and speciation. *Journal of Molecular Evolution*, *79*(1–2), 21–39. https://doi.org/10.1007/s00239-014-9633-0

Erbilgin, N., Ma, C., Whitehouse, C., Shan, B., Najar, A., & Evenden, M. (2014). Chemical similarity between historical and novel host plants promotes range and host expansion of the mountain pine beetle in a naïve host ecosystem. *New Phytologist*, *201*(3), 940–950. https://doi.org/10.1111/nph.12573

Espeland, M., Breinholt, J., Willmott, K. R., Warren, A. D., Vila, R., Toussaint, E. F. A., Maunsell, S. C., Aduse-Poku, K., Talavera, G., Eastwood, R., Jarzyna, M. A., Guralnick, R., Lohman, D. J., Pierce, N. E., & Kawahara, A. Y. (2018). A comprehensive and dated phylogenomic analysis of butterflies. *Current Biology*, *28*(5), 770--778.e5. https://doi.org/10.1016/j.cub.2018.01.061

Farkas, T. E., Hendry, A. P., Nosil, P., & Beckerman, A. P. (2015). How maladaptation can structure biodiversity: Eco-evolutionary island biogeography. *Trends in Ecology and Evolution*, *30*(3), 154–160. https://doi.org/10.1016/j.tree.2015.01.002

Farkas, T. E., Mononen, T., Comeault, A. A., & Nosil, P. (2016). Observational evidence that maladaptive gene flow reduces patch occupancy in a wild insect metapopulation. *Evolution*, *70*(12), 2879–2888. https://doi.org/10.1111/evo.13076

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Feyereisen, R. (1999). Insect P450 Enzymes. *Annual Review of Entomology*, *44*(1), 507–533. https://doi.org/10.1146/annurev.ento.44.1.507

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., & Punta, M. (2014). Pfam: The protein families database. *Nucleic Acids Research*, *42*(D1), D222–D230. https://doi.org/10.1093/nar/gkt1223

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford University Press.

Fitzpatrick, S. W., Gerberich, J. C., Kronenberger, J. A., Angeloni, L. M., & Funk, W. C. (2015). Locally adapted traits maintained in the face of high gene flow. *Ecology Letters*, *18*(1), 37–47. https://doi.org/10.1111/ELE.12388

Fitzpatrick, S. W., & Reid, B. N. (2019). Does gene flow aggravate or alleviate maladaptation to environmental stress in small populations? *Evolutionary Applications*, *12*(7), 1402–1416. https://doi.org/10.1111/eva.12768

Forde, S. E., Thompson, J. N., & Bohannan, B. J. M. (2004). Adaptation varies through space and time in a coevolving host-parasitoid interaction. *Nature*, *431*(7010), 841–844. https://doi.org/10.1038/nature02906

Forde, S. E., Thompson, J. N., & Bohannan, B. J. M. (2007). Gene flow reverses an adaptive cline in a coevolving host-parasitoid interaction. *American Naturalist*, *169*(6), 794–801. https://doi.org/10.1086/516848

Forsberg, J. (1987). Size discrimination among conspecific hostplants in two pierid butterflies; *Pieris napi* L. and *Pontia daplidice* L. *Oecologia*, *72*(1), 52–57. https://doi.org/10.1007/BF00385044

Franch-Gras, L., Hahn, C., García-Roger, E. M., Carmona, M. J., Serra, M., & Gómez, A. (2018). Genomic signatures of local adaptation to the degree of environmental predictability in rotifers. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-34188-y

Friberg, M., & Wiklund, C. (2019). Host preference variation cannot explain microhabitat differentiation among sympatric *Pieris napi* and *Pieris rapae* butterflies. *Ecological Entomology*, *44*(4), 571–576. https://doi.org/10.1111/een.12728

Futuyma, D. J., & Agrawal, A. A. (2009). Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences*, *106*(43), 18054–18061. https://doi.org/10.1073/pnas.0904106106

Gandon, S., Capowiez, Y., Dubois, Y., Michalakis, Y., & Olivieri, I. (1996). Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proceedings of the Royal Society B: Biological Sciences*, *263*(1373), 1003–1009. https://doi.org/10.1098/rspb.1996.0148

Gandon, S., & Nuismer, S. L. (2009). Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *American Naturalist*, *173*(2), 212–224. https://doi.org/10.1086/593706

Garant, D., Forde, S. E., & Hendry, A. P. (2007). The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, *21*(3), 434–443. https://doi.org/10.1111/j.1365-2435.2006.01228.x

García-Robledo, C., & Horvitz, C. C. (2012). Parent–offspring conflicts, “optimal bad motherhood” and the “mother knows best” principles in insect herbivores colonizing novel host plants. *Ecology and Evolution*, *2*(7), 1446–1457. https://doi.org/10.1002/ece3.267

Garud, N. R., Messer, P. W., & Petrov, D. A. (2021). Detection of hard and soft selective sweeps from *Drosophila melanogaster* population genomic data. *PLoS Genetics*, *17*(2 February), e1009373. https://doi.org/10.1371/journal.pgen.1009373

Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, *201*(4), 1555–1579. https://doi.org/10.1534/genetics.115.181453

Ge, S.-X., Shi, F.-M., Pei, J.-H., Hou, Z.-H., Zong, S.-X., & Ren, L.-L. (2021). Gut Bacteria Associated with *Monochamus saltuarius* (Coleoptera:

Cerambycidae) and Their Possible Roles in Host Plant Adaptations. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.687211

Geiger, H., & Shapiro, A. M. (1992). Genetics, systematics and evolution of holarctic *Pieris napi* species group populations (Lepidoptera, Pieridae). *Journal of Zoological Systematics and Evolutionary Research*, *30*(2), 100–122. https://doi.org/10.1111/j.1439-0469.1992.tb00161.x

Gering, E., Incorvaia, D., Henriksen, R., Wright, D., & Getty, T. (2019). Maladaptation in feral and domesticated animals. *Evolutionary Applications*, *12*(7), 1274–1286. https://doi.org/10.1111/eva.12784

Gilroy, J. J., & Sutherland, W. J. (2007). Beyond ecological traps: Perceptual errors and undervalued resources. *Trends in Ecology and Evolution*, *22*(7), 351–356. https://doi.org/10.1016/j.tree.2007.03.014

Gimbernat-Mayol, J., Mantes, A. D., Bustamante, C. D., Montserrat, D. M., & Ioannidis, A. G. (2022). Archetypal Analysis for population genetics. *PLOS Computational Biology*, *18*(8), e1010301. https://doi.org/10.1371/journal.pcbi.1010301

Girgis, H. Z. (2015). Red: An intelligent, rapid, accurate tool for detecting repeats de-novo on the genomic scale. *BMC Bioinformatics*, *16*(1), 227. https://doi.org/10.1186/s12859-015-0654-5

Gompert, Z., Saley, T., Philbin, C., Yoon, S. A., Perry, E., Sneck, M. E., Harrison, J. G., Buerkle, C. A., Fordyce, J. A., Nice, C. C., Dodson, C. D., Lebeis, S. L., Lucas, L. K., & Forister, M. L. (2022). Additive genetic effects in interacting species jointly determine the outcome of caterpillar herbivory. *Proceedings of the National Academy of Sciences*, *119*(36), e2206052119. https://doi.org/10.1073/pnas.2206052119

Gomulkiewicz, R., Drown, D. M., Dybdahl, M. F., Godsoe, W., Nuismer, S. L., Pepin, K. M., Ridenhour, B. J., Smith, C. I., & Yoder, J. B. (2007). Dos and don'ts of testing the geographic mosaic theory of coevolution. *Heredity*, *98*(5), 249–258. https://doi.org/10.1038/sj.hdy.6800949

Gomulkiewicz, R., & Holt, R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? *Evolution*, *49*(1), 201. https://doi.org/10.2307/2410305

Gomulkiewicz, R., Thompson, J. N., Holt, R. D., Nuismer, S. L., & Hochberg, M. E. (2000). Hot spots, cold spots, and the geographic mosaic theory of coevolution. *American Naturalist*, *156*(2), 156–174. https://doi.org/10.1086/303382

Gougherty, A. V., Keller, S. R., & Fitzpatrick, M. C. (2021). Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change*, 1–6. https://doi.org/10.1038/s41558-020-00968-6

Grant, P. R., & Grant, B. R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science*, *296*(5568), 707–711. https://doi.org/10.1126/science.1070315

Grant, P. R., & Grant, B. R. (2014). *40 Years of Evolution: Darwin's Finches on Daphne Major Island*. Princeton University Press. http://press.princeton.edu/titles/10282.html

Griffiths, J. S., Kawji, Y., & Kelly, M. W. (2021). An Experimental Test of Adaptive Introgression in Locally Adapted Populations of Splash Pool Copepods. *Molecular Biology and Evolution*, *38*(4), 1306–1316. https://doi.org/10.1093/molbev/msaa289

Gripenberg, S., Mayhew, P. J., Parnell, M., & Roslin, T. (2010). A meta-analysis of preference-performance relationships in phytophagous insects. *Ecology Letters*, *13*(3), 383–393. https://doi.org/10.1111/j.1461-0248.2009.01433.x

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220. https://doi.org/10.1534/genetics.113.152462

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genetics*, *5*(10), e1000695. https://doi.org/10.1371/journal.pgen.1000695

Hahn, D. A. (2005). Larval nutrition affects lipid storage and growth, but not protein or carbohydrate storage in newly eclosed adults of the grasshopper *Schistocerca americana*. *Journal of Insect Physiology*, *51*(11), 1210–1219. https://doi.org/10.1016/j.jinsphys.2005.06.011

Hahn, M. W., Han, M. V., & Han, S. G. (2007). Gene family evolution across 12 Drosophila genomes. *PLoS Genetics*, *3*(11), 2135–2146. https://doi.org/10.1371/JOURNAL.PGEN.0030197

Haldane, J. B. S. (1924). A mathematical theory of natural and artificial selection, part I. *Transactions of the Camrbidge Philosophical Society*, *23*, 19–41.

Haldane, J. B. S. (1932). *The causes of evolution*. London, New York Longmans, Green and Co.

Haldane, J. B. S. (1937). The Effect of Variation of Fitness. *The American Naturalist*, *71*(735), 337–349. https://doi.org/10.1086/280722

Hale, R., Morrongiello, J. R., & Swearer, S. E. (2016). Evolutionary traps and range shifts in a rapidly changing world. *Biology Letters*, *12*(6), 2016–2019. https://doi.org/10.1098/rsbl.2016.0003

Halkier, B. A., & Gershenzon, J. (2006). Biology and Biochemistry of Glucosinolates. *Annual Review of Plant Biology*, *57*(1), 303–333. https://doi.org/10.1146/annurev.arplant.57.032905.105228

Haller, B. C., & Hendry, A. P. (2014). Solving the paradox of stasis: Squashed stabilizing selection and the limits of detection. *Evolution*, *68*(2), 483–500. https://doi.org/10.1111/evo.12275

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLOS ONE*, *12*(10), e0185809. https://doi.org/10.1371/journal.pone.0185809

Halon, E., Eakteiman, G., Moshitzky, P., Elbaz, M., Alon, M., Pavlidi, N., Vontas, J., & Morin, S. (2015). Only a minority of broad-range detoxification genes respond to a variety of phytotoxins in generalist *Bemisia tabaci* species. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep17975

Hammer, T. J., & Bowers, M. D. (2015). Gut microbes may facilitate insect herbivory of chemically defended plants. *Oecologia*, *179*(1), 1–14. https://doi.org/10.1007/s00442-015-3327-1

Hansen, T. F., Carter, A. J. R., & Pélabon, C. (2006). On adaptive accuracy and precision in natural populations. *American Naturalist*, *168*(2), 168–181. https://doi.org/10.1086/505768

Hardy, N. B., & Otto, S. P. (2014). Specialization and generalization in the diversification of phytophagous insects: Tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1795), 20132960. https://doi.org/10.1098/rspb.2013.2960

Haribal, M., & Renwick, J. A. A. (1998). Isovitexin 6″-O-β-d-glucopyranoside: A feeding deterrent to *Pieris napi oleracea* from *Alliaria petiolata*. *Phytochemistry*, *47*(7), 1237–1240. https://doi.org/10.1016/S0031-9422(97)00740-1

Haribal, M., Yang, Z., Attygalle, A. B., Renwick, J. A. A., & Meinwald, J. (2001). A Cyanoallyl Glucoside from *Alliaria petiolata*, as a Feeding Deterrent for Larvae of *Pieris napi oleracea*. *Journal of Natural Products*, *64*(4), 440–443. https://doi.org/10.1021/np000534d

Haverkamp, A., Hansson, B. S., & Knaden, M. (2018). Combinatorial Codes and Labeled Lines: How Insects Use Olfactory Cues to Find and Judge Food,

Mates, and Oviposition Sites in Complex Environments. *Frontiers in Physiology*, *9*. https://www.frontiersin.org/articles/10.3389/fphys.2018.00049

Hedrick, P. W. (2011). *Genetics of populations* (4th ed). Jones and Bartlett Publishers.

Heidel-Fischer, H. M., & Vogel, H. (2015). Molecular mechanisms of insect adaptation to plant secondary compounds. *Current Opinion in Insect Science*, *8*, 8–14. https://doi.org/10.1016/j.cois.2015.02.004

Hendry, A. P., & Gonzalez, A. (2008). Whither adaptation? *Biology and Philosophy*, *23*(5), 673–699. https://doi.org/10.1007/s10539-008-9126-x

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, *173*(5), 579–588. https://doi.org/10.1086/597611

Hermisson, J., & Pennings, P. S. (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, *169*(4), 2335–2352. https://doi.org/10.1534/genetics.104.036947

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, *188*(4), 379–397. https://doi.org/10.1086/688018

Hoeksema, J. D., & Forde, S. E. (2008). A meta-analysis of factors affecting local adaptation between interacting species. *American Naturalist*, *171*(3), 275–290. https://doi.org/10.1086/527496

Hollings, T., Jones, M., Mooney, N., & Mccallum, H. (2014). Trophic Cascades Following the Disease-Induced Decline of an Apex Predator, the Tasmanian Devil. *Conservation Biology*, *28*(1), 63–75. https://doi.org/10.1111/cobi.12152

Holt, R. D. (1996). Adaptive Evolution in Source-Sink Environments: Direct and Indirect Effects of Density-Dependence on Niche Evolution. *Oikos*, *75*(2), 182. https://doi.org/10.2307/3546242

Hopkins, G. R., Gaston, K. J., Visser, M. E., Elgar, M. A., & Jones, T. M. (2018). Artificial light at night as a driver of evolution across urban–rural landscapes. *Frontiers in Ecology and the Environment*, *16*(8), 472–479. https://doi.org/10.1002/fee.1828

Horváth, G., Blahó, M., Egri, Á., Kriska, G., Seres, I., & Robertson, B. (2010). Reducing the maladaptive attractiveness of solar panels to polarotactic

insects. *Conservation Biology*, *24*(6), 1644–1653. https://doi.org/10.1111/j.1523-1739.2010.01518.x

Horváth, G., Malik, P., Kriska, G., & Wildermuth, H. (2007). Ecological traps for dragonflies in a cemetery: The attraction of *Sympetrum* species (Odonata: Libellulidae) by horizontally polarizing black gravestones. *Freshwater Biology*, *52*(9), 1700–1709. https://doi.org/10.1111/j.1365-2427.2007.01798.x

Huang, X. P., Renwick, J. A. A., & Chew, F. S. (1994). Oviposition stimulants and deterrents control acceptance of *Alliaria petiolata* by *Pieris rapae* and *P. napi oleracea*. *Chemoecology*, *5*(2), 79–87. https://doi.org/10.1007/BF01259436

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

Hufbauer, R. A., Szucs, M., Kasyon, E., Youngberg, C., Koontz, M. J., Richards, C., Tuff, T., & Melbourne, B. A. (2015). Three types of rescue can avert extinction in a changing environment. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10557–10562. https://doi.org/10.1073/pnas.1504732112

Ishihara, M., & Ohgushi, T. (2006). Reproductive Inactivity and Prolonged Developmental Time Induced by Seasonal Decline in Host Plant Quality in the Willow Leaf Beetle *Plagiodera versicolora* (Coleoptera: Chrysomelidae). *Environmental Entomology*, *35*(2), 524–530. https://doi.org/10.1603/0046-225X-35.2.524

Jaenike, J. (1978). On optimal oviposition behavior in phytophagous insects. *Theoretical Population Biology*, *14*(3), 350–356. https://doi.org/10.1016/0040-5809(78)90012-6

Jaenike, J. (1990). Host Specialization in Phytophagous Insects. *Annual Review of Ecology and Systematics*, *21*, 243–273.

Janz, N., & Nylin, S. (2008). *The oscillation hypothesis of host-plant range and speciation. Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects*. University of California Press Berkeley.

Janz, N., Nylin, S., & Wahlberg, N. (2006). Diversity begets diversity: Host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology*, *6*(1), 4. https://doi.org/10.1186/1471-2148-6-4

Janz, N., Nylin, S., & Wedell, N. (1994). Host plant utilization in the comma butterfly: Sources of variation and evolutionary implications. *Oecologia*, *99*(1), 132–140. https://doi.org/10.1007/BF00317093

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Jiang, X., Xu, Z., Li, J., Shi, Y., Wu, W., & Tao, S. (2011). The influence of deleterious mutations on adaptation in asexual populations. *PLoS ONE*, *6*(11), e27757. https://doi.org/10.1371/journal.pone.0027757

Jones, M. E., Jarman, P. J., Lees, C. M., Hesterman, H., Hamede, R. K., Mooney, N. J., Mann, D., Pukk, C. E., Bergfeld, J., & McCallum, H. (2007). Conservation management of Tasmanian devils in the context of an emerging, extinction-threatening disease: Devil facial tumor disease. *EcoHealth*, *4*(3), 326–337. https://doi.org/10.1007/s10393-007-0120-6

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Joshi, A., & Thompson, J. N. (1995). Trade-offs and the evolution of host specialization. *Evolutionary Ecology*, *9*(1), 82–92. https://doi.org/10.1007/BF01237699

Jump, A. S., & Woodward, F. I. (2003). Seed production and population density decline approaching the range-edge of *Cirsium* species. *New Phytologist*, *160*(2), 349–358. https://doi.org/10.1046/j.1469-8137.2003.00873.x

Kaltz, O., Gandon, S., Michalakis, Y., & Shykoff, J. A. (1999). Local maladaptation in the anther-smut fungus *Microbotryum violaceum* to its host plant <i>Silene latifolia</I>: Evidence from a cross-inoculation experiment. *Evolution*, *53*(2), 395–407. https://doi.org/10.1111/j.1558-5646.1999.tb03775.x

Kanarek, A. R., Webb, C. T., Barfield, M., & Holt, R. D. (2014). Overcoming allee effects through evolutionary, genetic, and demographic rescue. *Journal of Biological Dynamics*, *9*(1), 15–33. https://doi.org/10.1080/17513758.2014.978399

Kant, M. R., Jonckheere, W., Knegt, B., Lemos, F., Liu, J., Schimmel, B. C. J., Villarroel, C. A., Ataide, L. M. S., Dermauw, W., Glas, J. J., Egas, M., Janssen, A., Van Leeuwen, T., Schuurink, R. C., Sabelis, M. W., & Alba, J. M. (2015). Mechanisms and ecological consequences of plant defence induction and suppression in herbivore communities. *Annals of Botany*, *115*(7), 1015–1051. https://doi.org/10.1093/aob/mcv054

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Keeler, M. S., & Chew, F. S. (2008). Escaping an evolutionary trap: Preference and performance of a native insect on an exotic invasive host. *Oecologia*, *156*(3), 559–568. https://doi.org/10.1007/s00442-008-1005-2

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*(8), 782–788. https://doi.org/10.1111/2041-210X.12067

Kelkenberg, M., Odman-Naresh, J., Muthukrishnan, S., & Merzendorfer, H. (2015). Chitin is a necessary component to maintain the barrier function of the peritrophic matrix in the insect midgut. *Insect Biochemistry and Molecular Biology*, *56*, 21–28. https://doi.org/10.1016/j.ibmb.2014.11.005

Kelly, E., & Phillips, B. L. (2016). Targeted gene flow for conservation. *Conservation Biology*, *30*(2), 259–267. https://doi.org/10.1111/cobi.12623

King, K. C., Seppälä, O., & Neiman, M. (2012). Is more better? Polyploidy and parasite resistance. *Biology Letters*, *8*(4), 598–600. https://doi.org/10.1098/rsbl.2011.1152

King, R. B., & Lawson, R. (1995). Color-pattern variation in Lake Erie water snakes: The role of gene flow. *Evolution*, *49*(5), 885–896. https://doi.org/10.1111/j.1558-5646.1995.tb02324.x

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *The American Naturalist*, *150*(1), 1–23. https://doi.org/10.1086/286054

Kitahara, H. (2016). Oviposition plants and seasonal migratory movements of sympatric *Pieris melete* and *P. napi japonica* (Lepidoptera, Pieridae). *Lepidoptera Science*, *67*(1), 32–40.

Klemola, N., Klemola, T., Rantala, M. J., & Ruuhola, T. (2007). Natural host-plant quality affects immune defence of an insect herbivore. *Entomologia Experimentalis et Applicata*, *123*(2), 167–176. https://doi.org/10.1111/j.1570-7458.2007.00533.x

Kolde, R., & Kolde, M. R. (2018). Package 'pheatmap.' *R Package*, *1*(10).

Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, *21*(4), 1359–1368. https://doi.org/10.1111/1755-0998.13326

Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, *47*(D1), D807–D811. https://doi.org/10.1093/nar/gky1053

Krohn, A. R., Diepeveen, E. T., Bi, K., & Rosenblum, E. B. (2019). Local adaptation does not lead to genome-wide differentiation in lava flow lizards. *Ecology and Evolution*, *9*(12), 6810–6820. https://doi.org/10.1002/ECE3.5231

Krüger, O., Davies, N. B., & Sorenson, M. D. (2007). The evolution of sexual dimorphism in parasitic cuckoos: Sexual selection or coevolution? *Proceedings of the Royal Society B*, *274*(1617), 1553–1560.

Krüger, O., Lindström, J., & Amos, W. (2001). Maladaptive mate choice maintained by heterozygote advantage. *Evolution*, *55*(6), 1207–1214. https://doi.org/10.1111/j.0014-3820.2001.tb00640.x

Kuchernig, J. C., Burow, M., & Wittstock, U. (2012). Evolution of specifier proteins in glucosinolate-containing plants. *BMC Evolutionary Biology*, *12*(127). http://www.biomedcentral.com/1471-2148/12/127

Kudo, G., & Cooper, E. J. (2019). When spring ephemerals fail to meet pollinators: Mechanism of phenological mismatch and its impact on plant reproduction. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1904). https://doi.org/10.1098/rspb.2019.0573

Kudo, G., & Ida, T. Y. (2013). Early onset of spring increases the phenological mismatch between plants and pollinators. *Ecology*, *94*(10), 2311–2320. https://doi.org/10.1890/12-2003.1

Kuparinen, A., & Uusi-Heikkilä, S. (2020). Atlantic cod recovery from the Allee effect zone: Contrasting ecological and evolutionary rescue. *Fish and Fisheries*, *21*(5), 916–926. https://doi.org/10.1111/faf.12470

Kuussaari, M., Bommarco, R., Heikkinen, R. K., Helm, A., Krauss, J., Lindborg, R., Öckinger, E., Pärtel, M., Pino, J., Rodà, F., Stefanescu, C., Teder, T., Zobel, M., & Steffan-Dewenter, I. (2009). Extinction debt: A challenge for biodiversity conservation. *Trends in Ecology and Evolution*, *24*(10), 564–571. https://doi.org/10.1016/j.tree.2009.04.011

Ladner, D. T., & Altizer, S. (2005). Oviposition preference and larval performance of North American monarch butterflies on four *Asclepias* species. *Entomologia Experimentalis et Applicata*, *116*(1), 9–20. https://doi.org/10.1111/j.1570-7458.2005.00308.x

LaForgia, M. L., Harrison, S. P., & Latimer, A. M. (2020). Invasive species interact with climatic variability to reduce success of natives. *Ecology*, *101*(6), e03022. https://doi.org/10.1002/ecy.3022

Lande, R. (1975). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research*, *26*(3), 221–235. https://doi.org/10.1017/S0016672300016037

Lande, R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution*, *30*(2), 314–334. https://doi.org/10.2307/2407703

Landolt, P. J., & Phillips, T. W. (1997). Host Plant Influences on Sex Pheromone Behavior of Phytophagous Insects. *Annual Review of Entomology*, *42*(1), 371–391. https://doi.org/10.1146/annurev.ento.42.1.371

Lao, S., Robertson, B. A., Anderson, A. W., Blair, R. B., Eckles, J. W., Turner, R. J., & Loss, S. R. (2020). The influence of artificial night at night and polarized light on bird-building collisions. *Biological Conservation*, *241*, 108358. https://doi.org/10.1016/j.biocon.2019.108358

Larsson, S., & Ekbom, B. (1995). Oviposition Mistakes in Herbivorous Insects: Confusion or a Step Towards a New Host Plant? *Oikos*, *72*(1), 155–160. https://doi.org/10.2307/3546051

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, *17*(4), 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Leroy, T., Louvet, J.-M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J.-M., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist*, *226*(4), 1171–1182. https://doi.org/10.1111/nph.16095

Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM* (arXiv:1303.3997). arXiv. http://arxiv.org/abs/1303.3997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009a). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009b). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*(8), e47. https://doi.org/10.1093/nar/gkz114

Lindroth, R. L. (1989). Host plant alteration of detoxication activity in *Papilio glaucus glaucus*. *Entomologia Experimentalis et Applicata*, *50*(1), 29–35. https://doi.org/10.1111/j.1570-7458.1989.tb02310.x

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Liu, Y., Liu, H., Liu, S., Wang, S., Jiang, R.-J., & Li, S. (2009). Hormonal and nutritional regulation of insect fat body development and function. *Archives of Insect Biochemistry and Physiology*, *71*(1), 16–30. https://doi.org/10.1002/arch.20290

Loehlin, D. W., Oliveira, D. C. S. G., Edwards, R., Giebel, J. D., Clark, M. E., Cattani, M. V., van de Zande, L., Verhulst, E. C., Beukeboom, L. W., Muñoz-Torres, M., & Werren, J. H. (2010). Non-coding changes cause sex-specific wing size differences between closely related species of *Nasonia*. *PLoS Genetics*, *6*(1), e1000821. https://doi.org/10.1371/journal.pgen.1000821

Loewe, L., & Hill, W. G. (2010). The population genetics of mutations: Good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1544), 1153–1167. https://doi.org/10.1098/rstb.2009.0317

MacArthur, R. H. (1984). *Geographical ecology: Patterns in the distribution of species*. Princeton University Press.

MacPherson, A., Otto, S. P., & Nuismer, S. L. (2018). Keeping Pace with the Red Queen: Identifying the Genetic Basis of Susceptibility to Infectious Disease. *Genetics*, *208*(2), 779–789. https://doi.org/10.1534/GENETICS.117.300481

Magnan, A. K., Schipper, E. L. F., Burkett, M., Bharwani, S., Burton, I., Eriksen, S., Gemenne, F., Schaar, J., & Ziervogel, G. (2016). Addressing the risk of maladaptation to climate change. *Wiley Interdisciplinary Reviews: Climate Change*, *7*(5), 646–665. https://doi.org/10.1002/wcc.409

Mao, W., Berhow, M. A., Zangerl, A. R., Mcgovern, J., & Berenbaum, M. R. (2006). Cytochrome P450-Mediated Metabolism of Xanthotoxin by *Papilio multicaudatus*. *Journal of Chemical Ecology*, *32*(3), 523–536. https://doi.org/10.1007/s10886-005-9018-3

Märkle, H., John, S., Cornille, A., Fields, P. D., & Tellier, A. (2021). Novel genomic approaches to study antagonistic coevolution between hosts and parasites. *Molecular Ecology*, *30*(15), 3660–3676. https://doi.org/10.1111/MEC.16001

Martin, K., Stacey, P. B., & Braun, C. E. (2000). Recruitment, Dispersal, and Demographic Rescue in Spatially-Structured White-Tailed Ptarmigan Populations. *The Condor*, *102*(3), 503–516. https://doi.org/10.1093/condor/102.3.503

Mason, C. J., Ray, S., Shikano, I., Peiffer, M., Jones, A. G., Luthe, D. S., Hoover, K., & Felton, G. W. (2019). Plant defenses interact with insect enteric bacteria by initiating a leaky gut syndrome. *Proceedings of the National Academy of Sciences*, *116*(32), 15991–15996. https://doi.org/10.1073/pnas.1908748116

Matsunaga, T., Reisenman, C. E., Goldman-Huertas, B., Brand, P., Miao, K., Suzuki, H. C., Verster, K. I., Ramírez, S. R., & Whiteman, N. K. (2022). Evolution of Olfactory Receptors Tuned to Mustard Oils in Herbivorous Drosophilidae. *Molecular Biology and Evolution*, *39*(2), msab362. https://doi.org/10.1093/molbev/msab362

McIndoo, N. E. (1929). Tropisms and sense organs of Lepidoptera. *Smithsonian Miscellaneous Collections*.

Menacer, K., Cortesero, A. M., & Hervé, M. R. (2021). Challenging the Preference–Performance Hypothesis in an above-belowground insect. *Oecologia*, *197*(1), 179–187. https://doi.org/10.1007/s00442-021-05007-5

Mendes, F. K., Vanderpool, D., Fulton, B., & Hahn, M. W. (2020). CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics*, *36*(22–23), 5516–5518. https://doi.org/10.1093/bioinformatics/btaa1022

Missbach, C., Dweck, H. K., Vogel, H., Vilcinskas, A., Stensmyr, M. C., Hansson, B. S., & Grosse-Wilde, E. (2014). Evolution of insect olfactory receptors. *ELife*, *3*, e02115. https://doi.org/10.7554/eLife.02115

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/nar/gkaa913

Montejo-Kovacevich, G., Farkas, T., Beckerman, A., & Nosil, P. (2020). Exploring context dependency in eco-evolutionary patterns with the stick insect *Timema cristinae*. *Ecology and Evolution*, *10*(15), 8197–8209. https://doi.org/10.1002/ece3.6526

Morrissey, M. B., & De Kerckhove, D. T. (2009). The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *American Naturalist*, *174*(6), 875–889. https://doi.org/10.1086/648311

Muhlia-Almazán, A., Sánchez-Paz, A., & García-Carreño, F. L. (2008). Invertebrate trypsins: A review. *Journal of Comparative Physiology B*, *178*(6), 655–672. https://doi.org/10.1007/s00360-008-0263-y

Muller, H. J. (1950). Our load of mutations. *American Journal of Human Genetics*, *2*(2), 111–176.

Murad, N. F., Silva-Brandão, K. L., & Brandão, M. M. (2021). Mechanisms behind polyphagia in a pest insect: Responses of *Spodoptera frugiperda* (J.E. Smith) strains to preferential and alternative larval host plants assessed with gene regulatory networks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1864*(3), 194687. https://doi.org/10.1016/j.bbagrm.2021.194687

Nakajima, M. (2014). Dynamics and Structure of a Native Pieris Population in the Presence of a Non-Native, Toxic Larval Host Plant. *The Journal of the Lepidopterists' Society*, *68*(3), 175–184. https://doi.org/10.18473/lepi.v68i3.a4

Nakajima, M., & Boggs, C. L. (2015). Fine-grained distribution of a non-native resource can alter the population dynamics of a native consumer. *PLoS ONE*, *10*(11), 1–17. https://doi.org/10.1371/journal.pone.0143052

Nakajima, M., Boggs, C. L., Bailey, S., Reithel, J., & Paape, T. (2013). Fitness costs of butterfly oviposition on a lethal non-native plant in a mixed native and non-native plant community. *Oecologia*, *172*(3), 823–832.

Nallu, S., Hill, J. A., Don, K., Sahagun, C., Zhang, W., Meslin, C., Snell-Rood, E., Clark, N. L., Morehouse, N. I., Bergelson, J., Wheat, C. W., & Kronforst, M. R. (2018). The molecular genetic basis of herbivory between butterflies and their host plants. *Nature Ecology & Evolution*, *2*(9), Article 9. https://doi.org/10.1038/s41559-018-0629-9

Neafsey, D. E., Waterhouse, R. M., Abai, M. R., Aganezov, S. S., Alekseyev, M. A., Allen, J. E., Amon, J., Arcà, B., Arensburger, P., Artemov, G., Assour, L. A., Basseri, H., Berlin, A., Birren, B. W., Blandin, S. A., Brockman, A. I., Burkot, T. R., Burt, A., Chan, C. S., … Besansky, N. J. (2015). Highly evolvable malaria vectors: The genomes of 16 Anopheles mosquitoes. *Science*, *347*(6217), 1258522. https://doi.org/10.1126/science.1258522

Nei, M., Niimura, Y., & Nozawa, M. (2008). The evolution of animal chemosensory receptor gene repertoires: Roles of chance and necessity. *Nature Reviews Genetics*, *9*(12), Article 12. https://doi.org/10.1038/nrg2480

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

North, A., Pennanen, J., Ovaskainen, O., & Laine, A. L. (2011). Local adaptation in a changing world: The roles of gene-flow, mutation, and sexual reproduction. *Evolution*, *65*(1), 79–89. https://doi.org/10.1111/j.1558-5646.2010.01107.x

Nosil, P., Soria-Carrasco, V., Feder, J. L., Flaxman, S. M., & Gompert, Z. (2019). Local and system-wide adaptation is influenced by population connectivity. *Conservation Genetics*, *20*(1), 45–57. https://doi.org/10.1007/s10592-018-1097-0

Nuismer, S. L., & Thompson, J. N. (2001). Plant polyploidy and non-uniform effects on insect herbivores. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1479), 1937–1940. https://doi.org/10.1098/rspb.2001.1760

Nuismer, S. L., Thompson, J. N., & Gomulkiewicz, R. (2000). Coevolutionary clines across selection mosaics. *Evolution*, *54*(4), 1102–1115. https://doi.org/10.1111/j.0014-3820.2000.tb00546.x

Nuismer, S. L., Thompson, J. N., & Gomulkiewicz, R. (2003). Coevolution between hosts and parasites with partially overlapping geographic ranges. *Journal of Evolutionary Biology*, *16*(6), 1337–1345. https://doi.org/10.1046/j.1420-9101.2003.00609.x

Nylin, S., Bergström, A., & Janz, N. (2000). Butterfly Host Plant Choice in the Face of Possible Confusion. *Journal of Insect Behavior*, *13*(4), 469–482. https://doi.org/10.1023/A:1007839200323

Nylin, S., & Janz, N. (1993). Oviposition preference and larval performance in *Polygonia c-album* (Lepidoptera: Nymphalidae): The choice between bad and worse. *Ecological Entomology*, *18*(4), 394–398. https://doi.org/10.1111/j.1365-2311.1993.tb01116.x

Nylin, S., & Janz, N. (2009). Butterfly host plant range: An example of plasticity as a promoter of speciation? *Evolutionary Ecology*, *23*(1), 137–146. https://doi.org/10.1007/s10682-007-9205-5

Nylin, S., Slove, J., & Janz, N. (2014). Host Plant Utilization, Host Range Oscillations And Diversification In Nymphalid Butterflies: A Phylogenetic Investigation. *Evolution*, *68*(1), 105–124. https://doi.org/10.1111/evo.12227

Okamura, Y., Sato, A., Tsuzuki, N., Murakami, M., Heidel-Fischer, H., & Vogel, H. (2019). Molecular signatures of selection associated with host plant differences in *Pieris* butterflies. *Molecular Ecology*, *28*(22), 4958–4970. https://doi.org/10.1111/mec.15268

Okamura, Y., Sato, A., Tsuzuki, N., Sawada, Y., Hirai, M. Y., Heidel-Fischer, H., Reichelt, M., Murakami, M., & Vogel, H. (2019). Differential regulation of host plant adaptive genes in *Pieris* butterflies exposed to a range of glucosinolate profiles in their host plants. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-43703-8

Oppliger, A., Vernet, R., & Baez, M. (1999). Parasite local maladaptation in the Canarian lizard *Gallotia galloti* (Reptilia: Lacertidae) parasitized by haemogregarian blood parasite. *Journal of Evolutionary Biology*, *12*(5), 951–955. https://doi.org/10.1046/j.1420-9101.1999.00101.x

Orsucci, M., Audiot, P., Nidelet, S., Dorkeld, F., Pommier, A., Vabre, M., Severac, D., Rohmer, M., Gschloessl, B., & Streiff, R. (2018). Transcriptomic response of female adult moths to host and non-host plants in two closely related species. *BMC Evolutionary Biology*, *18*(1), 145. https://doi.org/10.1186/s12862-018-1257-3

Oswald, B. P., & Nuismer, S. L. (2007). Neopolyploidy and pathogen resistance. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1624), 2393–2397. https://doi.org/10.1098/rspb.2007.0692

Ozaki, K., Ryuda, M., Yamada, A., Utoguchi, A., Ishimoto, H., Calas, D., Marion-Poll, F., Tanimura, T., & Yoshikawa, H. (2011). A gustatory receptor involved in host plant recognition for oviposition of a swallowtail butterfly. *Nature Communications*, *2*(1), Article 1. https://doi.org/10.1038/ncomms1548

Pan, Y., Chai, P., Zheng, C., Xu, H., Wu, Y., Gao, X., Xi, J., & Shang, Q. (2018). Contribution of cytochrome P450 monooxygenase CYP380C6 to spirotetramat resistance in *Aphis gossypii* Glover. *Pesticide Biochemistry and Physiology*, *148*, 182–189. https://doi.org/10.1016/j.pestbp.2018.04.015

Pelletier, F., Garant, D., & Hendry, A. P. (2009). Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523), 1483–1489. https://doi.org/10.1098/rstb.2009.0027

Perrier, C., Rougemont, Q., & Charmantier, A. (2020). Demographic history and genomics of local adaptation in blue tit populations. *Evolutionary Applications*, *13*(6), 1145–1165. https://doi.org/10.1111/eva.13035

Petkova, D., Novembre, J., & Stephens, M. (2015). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. https://doi.org/10.1038/ng.3464

Picard toolkit. (2019). In *Broad Institute, GitHub repository*. Broad Institute. https://broadinstitute.github.io/picard/

Pimm, S. L., & Askins, R. A. (1995). Forest losses predict bird extinctions in eastern North America. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(20), 9343–9347. https://doi.org/10.1073/pnas.92.20.9343

Pivnick, K. A., Jarvis, B. J., & Slater, G. P. (1994). Identification of olfactory cues used in host-plant finding by diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *Journal of Chemical Ecology*, *20*(7), 1407–1427. https://doi.org/10.1007/BF02059870

Pomés, A., Melén, E., Vailes, L. D., Retief, J. D., Arruda, L. K., & Chapman, M. D. (1998). Novel Allergen Structures with Tandem Amino Acid Repeats Derived from German and American Cockroach. *Journal of Biological Chemistry*, *273*(46), 30801–30807. https://doi.org/10.1074/jbc.273.46.30801

Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: Hypothesis testing using phylogenies. *Bioinformatics (Oxford, England)*, *21*(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Prasai, K., & Karlsson, B. (2011). Variation in immune defence in relation to developmental pathway in the green-veined white butterfly, Pieris napi. *Evolutionary Ecology Research*, *13*(3), 295–305.

Pulliam, H. R. (1988). Sources, sinks and population regulation. *American Naturalist*, *132*(5), 652–661. https://doi.org/10.1086/284880

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

R Development Core Team. (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. url:

Radchuk, V., Reed, T., Teplitsky, C., van de Pol, M., Charmantier, A., Hassall, C., Adamík, P., Adriaensen, F., Ahola, M. P., Arcese, P., Miguel Avilés, J., Balbontin, J., Berg, K. S., Borras, A., Burthe, S., Clobert, J., Dehnhard, N., de Lope, F., Dhondt, A. A., … Kramer-Schadt, S. (2019). Adaptive responses of animals to climate change are most likely insufficient. *Nature Communications*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-10924-4

Ramaswamy, S. B., Ma, W. K., & Baker, G. T. (1987). Sensory cues and receptors for oviposition by *Heliothis virescens*. *Entomologia Experimentalis et Applicata*, *43*(2), 159–168. https://doi.org/10.1111/j.1570-7458.1987.tb03600.x

Rane, R. V., Ghodke, A. B., Hoffmann, A. A., Edwards, O. R., Walsh, T. K., & Oakeshott, J. G. (2019). Detoxifying enzyme complements and host use phenotypes in 160 insect species. *Current Opinion in Insect Science*, *31*, 131–138. https://doi.org/10.1016/j.cois.2018.12.008

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H., & Feyereisen, R. (2002). Evolution of supergene families associated with insecticide resistance. *Science*, *298*(5591), 179–181. https://doi.org/10.1126/science.1076781

Ranz, J. M., González, P. M., Clifton, B. D., Nazario-Yepiz, N. O., Hernández-Cervantes, P. L., Palma-Martínez, M. J., Valdivia, D. I., Jiménez-Kaufman, A., Lu, M. M., Markow, T. A., & Abreu-Goodger, C. (2021). A de novo transcriptional atlas in *Danaus plexippus* reveals variability in dosage compensation across tissues. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-021-02335-3

Ratzka, A., Vogel, H., Kliebenstein, D. J., Mitchell-Olds, T., & Kroymann, J. (2002). Disarming the mustard oil bomb. *Proceedings of the National Academy of Sciences*, *99*(17), 11223–11228. https://doi.org/10.1073/pnas.172112899

Rausher, M. D. (1979). Larval Habitat Suitability and Oviposition Preference in Three Related Butterflies. *Ecology*, *60*(3), 503–511. https://doi.org/10.2307/1936070

Rawlings, N., & Barrett, A. (1995). *Evolutionary families of metallopeptidases. 574 Methods Enzymol*.

Reed, D. H., Lowe, E. H., Briscoe, D. A., & Frankham, R. (2003). Fitness and Adaptation in a Novel Environment: Effect of Inbreeding, Prior

Environment, and Lineage. *Evolution*, *57*(8), 1822–1828. https://doi.org/10.1111/j.0014-3820.2003.tb00589.x

Rendón-Anaya, M., Wilson, J., Sveinsson, S., Fedorkov, A., Cottrell, J., Bailey, M. E. S., Ruņģis, D., Lexer, C., Jansson, S., Robinson, K. M., Street, N. R., & Ingvarsson, P. K. (2021). Adaptive Introgression Facilitates Adaptation to High Latitudes in European Aspen (*Populus tremula* L.). *Molecular Biology and Evolution*, *38*(11), 5034–5050. https://doi.org/10.1093/molbev/msab229

Renwick, J. A. A., & Chew, F. S. (1994). Oviposition Behavior in Lepidoptera. *Annual Review of Entomology*, *39*(1), 377–400. https://doi.org/10.1146/annurev.en.39.010194.002113

Reznick, D. (2016). Hard and Soft Selection Revisited: How Evolution by Natural Selection Works in the Real World. *Journal of Heredity*, *107*(1), 3–14. https://doi.org/10.1093/jhered/esv076

Reznick, D. N., & Ghalambor, C. K. (2001). The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution. In *Microevolution Rate, Pattern, Process* (pp. 183–198). Springer, Dordrecht. https://doi.org/10.1007/978-94-010-0585-2_12

Richard, M., Tallamy, D. W., & Mitchell, A. B. (2019). Introduced plants reduce species interactions. *Biological Invasions*, *21*(3), 983–992. https://doi.org/10.1007/s10530-018-1876-z

Robertson, B. A., & Blumstein, D. T. (2019). How to disarm an evolutionary trap. *Conservation Science and Practice*, *1*(11), e116. https://doi.org/10.1111/csp2.116

Robertson, B. A., & Chalfoun, A. D. (2016). Evolutionary traps as keys to understanding behavioral maladapation. *Current Opinion in Behavioral Sciences*, *12*, 12–17. https://doi.org/10.1016/j.cobeha.2016.08.007

Robertson, B. A., & Horváth, G. (2019). Color polarization vision mediates the strength of an evolutionary trap. *Evolutionary Applications*, *12*(2), 175–186. https://doi.org/10.1111/eva.12690

Robertson, B. A., & Hutto, R. L. (2006). A framework for understanding ecological traps and an evaluation of existing evidence. *Ecology*, *87*(5), 1075–1085. https://doi.org/10.1890/0012-9658(2006)87[1075:AFFUET]2.0.CO;2

Robertson, B. A., Rehage, J. S., & Sih, A. (2013). Ecological novelty and the emergence of evolutionary traps. *Trends in Ecology and Evolution*, *28*(9), 552–560. https://doi.org/10.1016/j.tree.2013.04.004

Robinson, G. S., Ackery, P. R., Kitching, I. J., Beccaloni, G. W., & Hernández, L. M. (2010). *HOSTS - A Database of the World's Lepidopteran Hostplants*. Natural History Museum. https://www.nhm.ac.uk/our-science/data/hostplants/index.html

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodman, J. E., & Chew, F. S. (1980). Phytochemical correlates of herbivory in a community of native and naturalized Cruciferae. *Biochemical Systematics and Ecology*, *8*(1), 43–50. https://doi.org/10.1016/0305-1978(80)90024-1

Román-Palacios, C., & Wiens, J. J. (2020). Recent responses to climate change reveal the drivers of species extinction and survival. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(8), 4211–4217. https://doi.org/10.1073/pnas.1913007117

Rytz, R., Croset, V., & Benton, R. (2013). Ionotropic Receptors (IRs): Chemosensory ionotropic glutamate receptors in Drosophila and beyond. *Insect Biochemistry and Molecular Biology*, *43*(9), 888–897. https://doi.org/10.1016/j.ibmb.2013.02.007

Ryuda, M., Calas-List, D., Yamada, A., Marion-Poll, F., Yoshikawa, H., Tanimura, T., & Ozaki, K. (2013). Gustatory Sensing Mechanism Coding for Multiple Oviposition Stimulants in the Swallowtail Butterfly, *Papilio xuthus*. *Journal of Neuroscience*, *33*(3), 914–924. https://doi.org/10.1523/JNEUROSCI.1405-12.2013

Sánchez-Gracia, A., Vieira, F. G., & Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity*, *103*(3), 208–216. https://doi.org/10.1038/hdy.2009.55

Saul, W.-C., & Jeschke, J. M. (2015). Eco-evolutionary experience in novel species interactions. *Ecology Letters*, *18*(3), 236–245. https://doi.org/10.1111/ele.12408

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807–820. https://doi.org/10.1038/nrg3522

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 595–619. https://doi.org/10.1146/annurev.ecolsys.38.091206.095646

Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., … Sherry, S. T. (2021). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *49*(D1), D10–D17. https://doi.org/10.1093/nar/gkaa892

Schlaepfer, M. A., Runge, M. C., & Sherman, P. W. (2002). Ecological and evolutionary traps. *Trends in Ecology and Evolution*, *17*(10), 474–480. https://doi.org/10.1016/S0169-5347(02)02580-6

Schlaepfer, M. A., Sherman, P. W., Blossey, B., & Runge, M. C. (2005). Introduced species as evolutionary traps. *Ecology Letters*, *8*(3), 241–246. https://doi.org/10.1111/j.1461-0248.2005.00730.x

Schluter, D., Marchinko, K. B., Barrett, R. D. H., & Rogers, S. M. (2010). Natural selection and the genetics of adaptation in threespine stickleback. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1552), 2479–2486. https://doi.org/10.1098/rstb.2010.0036

Schweiger, O., Biesmeijer, J. C., Bommarco, R., Hickler, T., Hulme, P. E., Klotz, S., Kühn, I., Moora, M., Nielsen, A., Ohlemüller, R., Petanidou, T., Potts, S. G., Pyšek, P., Stout, J. C., Sykes, M. T., Tscheulin, T., Vilà, M., Walther, G.-R., Westphal, C., … Settele, J. (2010). Multiple stressors on biotic interactions: How climate change and alien species interact to affect pollination. *Biological Reviews*, *85*(4), 777–795. https://doi.org/10.1111/j.1469-185X.2010.00125.x

Sellis, D., Callahan, B. J., Petrov, D. A., & Messer, P. W. (2011). Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(51), 20666–20671. https://doi.org/10.1073/pnas.1114573108

Semper-Pascual, A., Decarre, J., Baumann, M., Busso, J. M., Camino, M., Gómez-Valencia, B., & Kuemmerle, T. (2019). Biodiversity loss in deforestation frontiers: Linking occupancy modelling and physiological stress indicators to understand local extinctions. *Biological Conservation*, *236*, 281–288. https://doi.org/10.1016/j.biocon.2019.05.050

Senner, N. R., Stager, M., & Sandercock, B. K. (2017). Ecological mismatches are moderated by local conditions for two populations of a long-distance migratory bird. *Oikos*, *126*(1), 61–72. https://doi.org/10.1111/oik.03325

Seppey, M., Ioannidis, P., Emerson, B. C., Pitteloud, C., Robinson-Rechavi, M., Roux, J., Escalona, H. E., McKenna, D. D., Misof, B., Shin, S., Zhou, X., Waterhouse, R. M., & Alvarez, N. (2019). Genomic signatures

accompanying the dietary shift to phytophagy in polyphagan beetles. *Genome Biology*, *20*(1), 98. https://doi.org/10.1186/s13059-019-1704-5

Shafer, A. B. a, & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters*, *16*(7), 940–950. https://doi.org/10.1111/ele.12120

Shimada, Y., Shikano, T., & Merilä, J. (2011). A high incidence of selection on physiologically important genes in the three-spined stickleback, *Gasterosteus aculeatus*. *Molecular Biology and Evolution*, *28*(1), 181–193. https://doi.org/10.1093/molbev/msq181

Siemens, D. H., & Roy, B. A. (2005). Tests for parasite-mediated frequency-dependent selection in natural populations of an asexual plant species. *Evolutionary Ecology*, *19*(4), 321–338. https://doi.org/10.1007/s10682-005-6639-5

Siepielski, A. M., Morrissey, M. B., Buoro, M., Carlson, S. M., Caruso, C. M., Clegg, S. M., Coulson, T., DiBattista, J., Gotanda, K. M., Francis, C. D., Hereford, J., Kingsolver, J. G., Augustine, K. E., Kruuk, L. E. B., Martin, R. A., Sheldon, B. C., Sletvold, N., Svensson, E. I., Wade, M. J., & MacColl, A. D. C. (2017). Precipitation drives global variation in natural selection. *Science*, *355*(6328), 959–962. https://doi.org/10.1126/science.aag2773

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simons, A. M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1712), 1601–1609. https://doi.org/10.1098/rspb.2011.0176

Sinervo, B., Méndez-de-la-Cruz, F., Miles, D. B., Heulin, B., Bastiaans, E., Cruz, M. V. S., Lara-Resendiz, R., Martínez-Méndez, N., Calderón-Espinosa, M. L., Meza-Lázaro, R. N., Gadsden, H., Avila, L. J., Morando, M., De La Riva, I. J., Sepúlveda, P. V., Rocha, C. F. D., Ibargüengoytía, N., Puntriano, C. A., Massot, M., … Sites, J. W. (2010). Erosion of lizard diversity by climate change and altered thermal niches. *Science*, *328*(5980), 894–899. https://doi.org/10.1126/science.1184695

Singer, M. C. (2021). Preference Provides a Plethora of Problems (Don't Panic). *Annual Review of Entomology*, *66*, 1–22. https://doi.org/10.1146/annurev-ento-022720-061725

Singer, M. C., & Parmesan, C. (2010). Phenological asynchrony between herbivorous insects and their hosts: Signal of climate change or pre-

existing adaptive strategy? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3161–3176. https://doi.org/10.1098/rstb.2010.0144

Singer, M. C., & Parmesan, C. (2018). Lethal trap created by adaptive evolutionary response to an exotic resource. *Nature*, *557*(7704), 238–241. https://doi.org/10.1038/s41586-018-0074-6

Singer, M. C., & Parmesan, C. (2019a). Butterflies embrace maladaptation and raise fitness in colonizing novel host. *Evolutionary Applications*, *12*(7), 1417–1433. https://doi.org/10.1111/eva.12775

Singer, M. C., & Parmesan, C. (2019b). Butterflies embrace maladaptation and raise fitness in colonizing novel host. *Evolutionary Applications*, *12*(7), 1417–1433. https://doi.org/10.1111/eva.12775

Spieth, H. R., & Cordes, R. (2012). Geographic comparison of seasonal migration events of the large white butterfly, *Pieris brassicae*. *Ecological Entomology*, *37*(6), 439–445. https://doi.org/10.1111/j.1365-2311.2012.01385.x

Sprouffske, K., Aguílar-Rodríguez, J., Sniegowski, P., & Wagner, A. (2018). High mutation rates limit evolutionary adaptation in *Escherichia coli*. *PLoS Genetics*, *14*(4), e1007324. https://doi.org/10.1371/journal.pgen.1007324

Stapley, J., Reger, J., Feulner, P. G. D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., & Slate, J. (2010). Adaptation genomics: The next generation. *Trends in Ecology and Evolution*, *25*(12), 705–712. https://doi.org/10.1016/j.tree.2010.09.002

Stefanescu, C., Jubany, J., & Dantart, J. (2012). Egg–laying by the butterfly *Iphiclides podalirius* (Lepidoptera, Papilionidae) on alien plants: A broadening of host range or oviposition mistakes? *Animal Biodiversity and Conservation*, *29*(1), 83–90.

Steward, R. A., & Boggs, C. L. (2020). Experience may outweigh cue similarity in maintaining a persistent host-plant-based evolutionary trap. *Ecological Monographs*, *90*(3), e01412. https://doi.org/10.1002/ecm.1412

Steward, R. A., Epanchin-Niell, R. S., & Boggs, C. L. (2022). Novel host unmasks heritable variation in plant preference within an insect population. *Evolution*. https://doi.org/10.1111/EVO.14608

Steward, R. A., Fisher, L. M., & Boggs, C. L. (2019a). Pre- and post-ingestive defenses affect larval feeding on a lethal invasive host plant. *Entomologia Experimentalis et Applicata*, *167*(4), 292–305. https://doi.org/10.1111/EEA.12773

Steward, R. A., Fisher, L. M., & Boggs, C. L. (2019b). Pre- and post-ingestive defenses affect larval feeding on a lethal invasive host plant. *Entomologia Experimentalis et Applicata*, *167*(4), 292–305. https://doi.org/10.1111/eea.12773

Steward, R. A., Okamura, Y., Boggs, C. L., Vogel, H., & Wheat, C. W. (2021). The Genome of the Margined White Butterfly (*Pieris macdunnoughii*): Sex Chromosome Insights and the Power of Polishing with PoolSeq Data. *Genome Biology and Evolution*, *13*(4). https://doi.org/10.1093/GBE/EVAB053

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE*, *6*(7), e21800. https://doi.org/10.1371/journal.pone.0021800

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, *34*(suppl_2), W609–W612. https://doi.org/10.1093/nar/gkl315

Suzuki, H. C., Ozaki, K., Makino, T., Uchiyama, H., Yajima, S., & Kawata, M. (2018). Evolution of Gustatory Receptor Gene Family Provides Insights into Adaptation to Diverse Host Plants in Nymphalid Butterflies. *Genome Biology and Evolution*, *10*(6), 1351–1362. https://doi.org/10.1093/GBE/EVY093

Svensson, E. I., & Connallon, T. (2019). How frequency-dependent selection affects population fitness, maladaptation and evolutionary rescue. *Evolutionary Applications*, *12*(7), 1243–1258. https://doi.org/10.1111/eva.12714

Swift, R. J., Rodewald, A. D., & Senner, N. R. (2017). Breeding habitat of a declining shorebird in a changing environment. *Polar Biology*, *40*(9), 1777–1786. https://doi.org/10.1007/s00300-017-2101-8

Szaz, D., Horvath, G., Barta, A., Robertson, B. A., Farkas, A., Egri, A., Tarjanyi, N., Racz, G., & Kriska, G. (2015). Lamp-lit bridges as dual light-traps for the night-swarming mayfly, *Ephoron virgo*: Interaction of polarized and unpolarized light pollution. *PLoS ONE*, *10*(3), e0121194. https://doi.org/10.1371/journal.pone.0121194

Takami, Y., Koshio, C., Ishii, M., Fujii, H., Hidaka, T., & Shimizu, I. (2004). Genetic diversity and structure of urban populations of Pieris butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, *13*(2), 245–258.

Tephly, T. R., & Burchell, B. (1990). UDP-glucuronosyltransferases: A family of detoxifying enzymes. *Trends in Pharmacological Sciences*, *11*(7), 276–279. https://doi.org/10.1016/0165-6147(90)90008-V

Thompson, J. N. (1999a). Specific hypotheses on the geographic mosaic of coevolution. *American Naturalist*, *153*(SUPPL.). https://doi.org/10.1086/303208

Thompson, J. N. (1999b). The evolution of species interactions. *Science*, *284*(5423), 2116–2118. https://doi.org/10.1126/science.284.5423.2116

Thompson, J. N. (2005). The Geographic Mosaic of Coevolution. In *The Geographic Mosaic of Coevolution*. University of Chicago Press. https://doi.org/10.7208/9780226118697

Thompson, J. N., Nuismer, S. L., & Gomulkiewicz, R. (2002). Coevolution and Maladaptation. *Integrative and Comparative Biology*, *42*(2), 381–387. https://doi.org/10.1093/icb/42.2.381

Thompson, J. N., & Pellmyr, O. (1991). Evolution of Oviposition Behavior and Host Preference in Lepidoptera. *Annual Review of Entomology*, *36*(1), 65–89. https://doi.org/10.1146/annurev.en.36.010191.000433

Thornton, K. R., Jensen, J. D., Becquet, C., & Andolfatto, P. (2007). Progress and prospects in mapping recent selection in the genome. *Heredity*, *98*(6), 340–348. https://doi.org/10.1038/sj.hdy.6800967

Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, *25*(10), 2144–2164. https://doi.org/10.1111/mec.13606

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), Article 3. https://doi.org/10.1038/nprot.2012.016

Trowbridge, C. D., & Todd, C. D. (2001). Host-Plant Change in Marine Specialist Herbivores: Ascoglossan Sea Slugs on Introduced Macroalgae. *Ecological Monographs*, *71*(2), 219–243.

Tufto, J. (2001). Effects of releasing maladapted individuals: A demographic-evolutionary model. *American Naturalist*, *158*(4), 331–340. https://doi.org/10.1086/321987

Urban, M. C. (2006). Maladaptation and mass effects in a metacommunity: Consequences for species coexistence. *American Naturalist*, *168*(1), 28–40. https://doi.org/10.1086/505159

Urban, M. C. (2011). The evolution of species interactions across natural landscapes. *Ecology Letters*, *14*(7), 723–732. https://doi.org/10.1111/j.1461-0248.2011.01632.x

Urban, M. C., & Skelly, D. K. (2006). Evolving metacommunities: Toward an evolutionary perspective on metacommunities. *Ecology*, *87*(7), 1616–1626.

Valiente-Banuet, A., Aizen, M. A., Alcántara, J. M., Arroyo, J., Cocucci, A., Galetti, M., García, M. B., García, D., Gómez, J. M., Jordano, P., Medel, R., Navarro, L., Obeso, J. R., Oviedo, R., Ramírez, N., Rey, P. J., Traveset, A., Verdú, M., & Zamora, R. (2015). Beyond species loss: The extinction of ecological interactions in a changing world. *Functional Ecology*, *29*(3), 299–307. https://doi.org/10.1111/1365-2435.12356

Valladares, G., & Lawton, J. H. (1991). Host-Plant Selection in the Holly Leaf-Miner: Does Mother Know Best? *Journal of Animal Ecology*, *60*(1), 227–240. https://doi.org/10.2307/5456

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Vermeer, K. M. C. A., Dicke, M., & de Jong, P. W. (2011). The potential of a population genomics approach to analyse geographic mosaics of plant—Insect coevolution. *Evolutionary Ecology*, *25*(5), 977–992. https://doi.org/10.1007/s10682-010-9452-8

Vermeij, G. J., & Vermeij, A. (1974). Adaptation, Versatility, and Evolution. In *Syst. Zool* (Vol. 22, pp. 466–477). https://academic.oup.com/sysbio/article/22/4/466/1664729

Vertacnik, K. L., & Linnen, C. R. (2017). Evolutionary genetics of host shifts in herbivorous insects: Insights from the age of genomics. *Annals of the New York Academy of Sciences*, *1389*(1), 186–212. https://doi.org/10.1111/nyas.13311

Vilà-Cabrera, A., Premoli, A. C., & Jump, A. S. (2019). Refining predictions of population decline at species' rear edges. *Global Change Biology*, *25*(5), 1549–1560. https://doi.org/10.1111/gcb.14597

Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., & Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*, *118*(2), e2023989118. https://doi.org/10.1073/pnas.2023989118

Wang, H., Shi, Y., Wang, L., Liu, S., Wu, S., Yang, Y., Feyereisen, R., & Wu, Y. (2018). CYP6AE gene cluster knockout in *Helicoverpa armigera* reveals role in detoxification of phytochemicals and insecticides. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-07226-6

Warwick, S. I., Francis, A., & Susko, D. J. (2011). The biology of Canadian weeds. 9. *Thlaspi arvense* L. (updated). *Https://Doi.Org/10.4141/P01-159*, *82*(4), 803–823. https://doi.org/10.4141/P01-159

Watt, W. B., Wheat, C. W., Meyer, E. H., & Martin, J. F. (2003). Adaptation at specific loci. VII. Natural selection, dispersal and the diversity of molecular-functional variation patterns among butterfly species complexes (*Colias*: Lepidoptera, Pieridae). *Molecular Ecology*, *12*(5), 1265–1275. https://doi.org/10.1046/j.1365-294X.2003.01804.x

Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., & Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution*, *35*(3), 773–777. https://doi.org/10.1093/molbev/msx335

Wheat, C. W., Vogel, H., Wittstock, U., Braby, M. F., Underwood, D., & Mitchell-Olds, T. (2007). The genetic basis of a plant–insect coevolutionary key innovation. *Proceedings of the National Academy of Sciences*, *104*(51), 20427–20431. https://doi.org/10.1073/pnas.0706229104

Whitlock, M. C. (2015). Modern approaches to local adaptation. *American Naturalist*, *186*, S1–S4.

Wiens, J. J., Lapoint, R. T., & Whiteman, N. K. (2015). Herbivory increases diversification across insect clades. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms9370

Wiklund, C. (1975). The evolutionary relationship between adult oviposition preferences and larval host plant range in *Papilio machaon* L. *Oecologia*, *18*(3), 185–197. https://doi.org/10.1007/BF00345421

Willi, Y., Fracassetti, M., Zoller, S., & Van Buskirk, J. (2018). Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution*, *35*(4), 781–791. https://doi.org/10.1093/molbev/msy003

Williams, G. C. (1966). Natural Selection, the Costs of Reproduction, and a Refinement of Lack's Principle. *The American Naturalist*, *100*(916), 687–690. https://doi.org/10.1086/282461

Winchell, K. M., Reynolds, R. G., Prado-Irwin, S. R., Puente-Rolón, A. R., & Revell, L. J. (2016). Phenotypic shifts in urban areas in the tropical lizard

Anolis cristatellus. *Evolution*, *70*(5), 1009–1022. https://doi.org/10.1111/evo.12925

Winfree, R., Aguilar, R., Vázquez, D. P., Lebuhn, G., & Aizen, M. A. (2009). A meta-analysis of bees' responses to anthropogenic disturbance. *Ecology*, *90*(8), 2068–2076. https://doi.org/10.1890/08-1245.1

Wissinger, S. A., & Whiteman, H. H. (1992). Fluctuation in a Rocky Mountain population of salamanders: Anthropogenic acidification or natural variation? *Journal of Herpetology*, *26*(4), 377–391. https://doi.org/10.2307/1565114

Witherington, B. E., & Bjorndal, K. A. (1991). Influences of artificial lighting on the seaward orientation of hatchling loggerhead turtles *Caretta caretta*. *Biological Conservation*, *55*(2), 139–149. https://doi.org/10.1016/0006-3207(91)90053-C

Wittstock, U., & Halkier, B. A. (2002). Glucosinolate research in the Arabidopsis era. *Trends in Plant Science*, *7*(6), 263–270. https://doi.org/10.1016/S1360-1385(02)02273-2

Xu, W. (2020). How do moth and butterfly taste?—Molecular basis of gustatory receptors in Lepidoptera. *Insect Science*, *27*(6), 1148–1157. https://doi.org/10.1111/1744-7917.12718

Yang, K., Gong, X.-L., Li, G.-C., Huang, L.-Q., Ning, C., & Wang, C.-Z. (2020). A gustatory receptor tuned to the steroid plant hormone brassinolide in *Plutella xylostella* (Lepidoptera: Plutellidae). *ELife*, *9*, e64114. https://doi.org/10.7554/eLife.64114

Yeaman, S., & Otto, S. P. (2011). Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution*, *65*(7), 2123–2129. https://doi.org/10.1111/j.1558-5646.2011.01277.x

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution*, *65*(7), 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Yoon, S., & Read, Q. (2016). Consequences of exotic host use: Impacts on Lepidoptera and a test of the ecological trap hypothesis. *Oecologia*, *181*(4), 985–996. https://doi.org/10.1007/s00442-016-3560-2

You, Y., Xie, M., Ren, N., Cheng, X., Li, J., Ma, X., Zou, M., Vasseur, L., Gurr, G. M., & You, M. (2015). Characterization and expression profiling of glutathione S-transferases in the diamondback moth, *Plutella xylostella* (L.). *BMC Genomics*, *16*(1), 152. https://doi.org/10.1186/s12864-015-1343-5

Appendix A:

CHAPTER 2 SUPPLEMENTARY INFORMATION

A.1 Tables

Table A.1. Proposal variances used for EEMS runs.

Run Parameter	400 demes, run 1	400 demes, run 2	800 demes, run 1	800 demes, run 2
numMCMCIter	15000000	15000000	15000000	15000000
numBurnIter	1500000	1500000	1500000	1500000
numThinIter	9999	9999	9999	9999
negBiSize	10	10	10	10
negBiProb	0.670000	0.670000	0.670000	0.670000
qVoronoiPr	0.250000	0.250000	0.250000	0.250000
mrateShape	0.000500	0.000500	0.000500	0.000500
qrateShape	0.002000	0.002000	0.002000	0.002000
sigmaShape	0.001000	0.001000	0.001000	0.001000
qrateScale	0.500000	0.500000	0.500000	0.500000
mrateScale	2.000000	2.000000	2.000000	2.000000
sigmaScale	1.000000	1.000000	1.000000	1.000000
mSeedsProposalS2	0.000290	0.000290	0.000250	0.000250
qSeedsProposalS2	0.002000	0.002000	0.002600	0.002600
mEffctProposalS2	0.520000	0.520000	0.515000	0.515000
qEffctProposalS2	0.007100	0.007100	0.010000	0.010000
mrateMuProposalS2	0.015000	0.015000	0.015000	0.015000

A.2 Figures

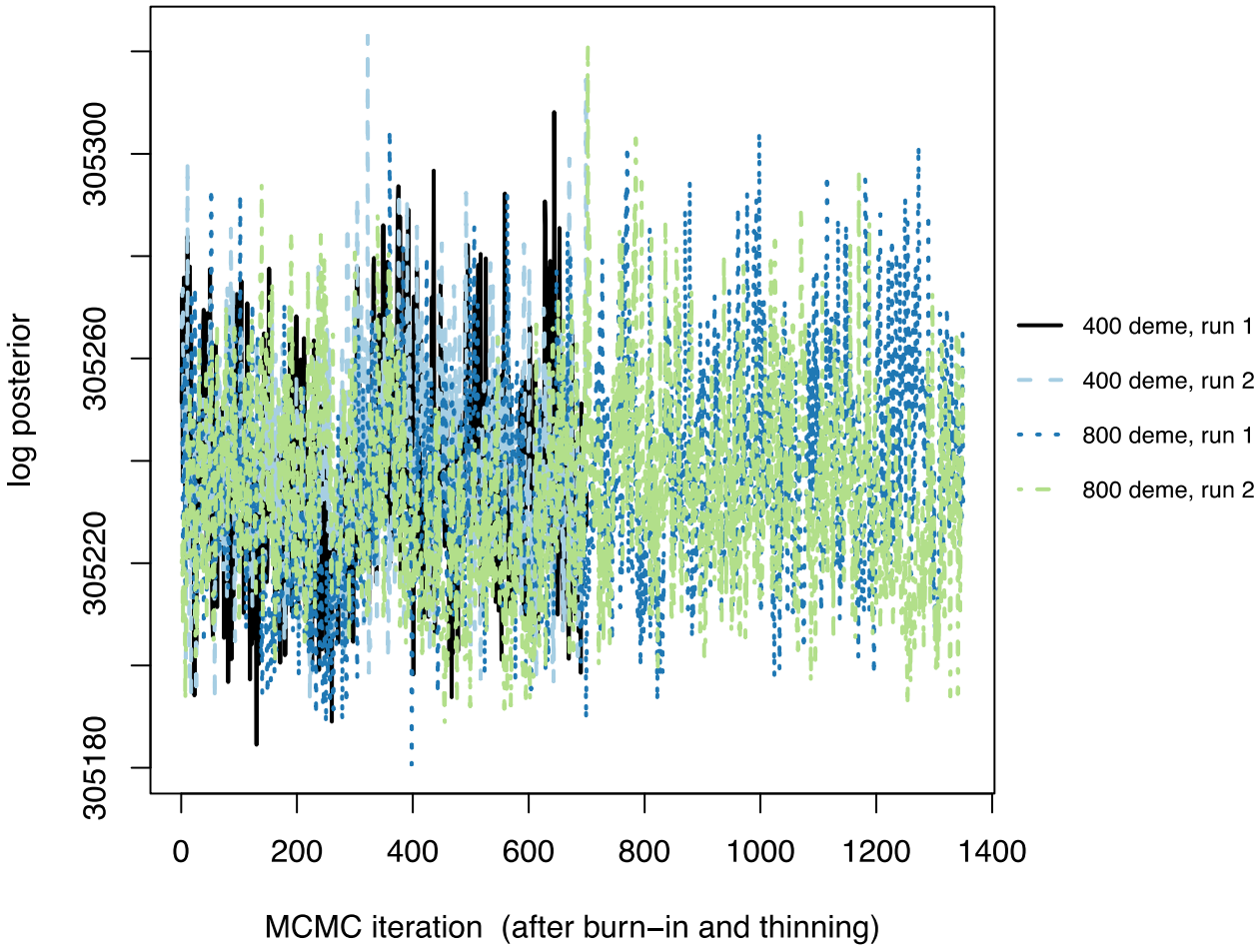

Figure A.1. Model convergence from EEMS runs.

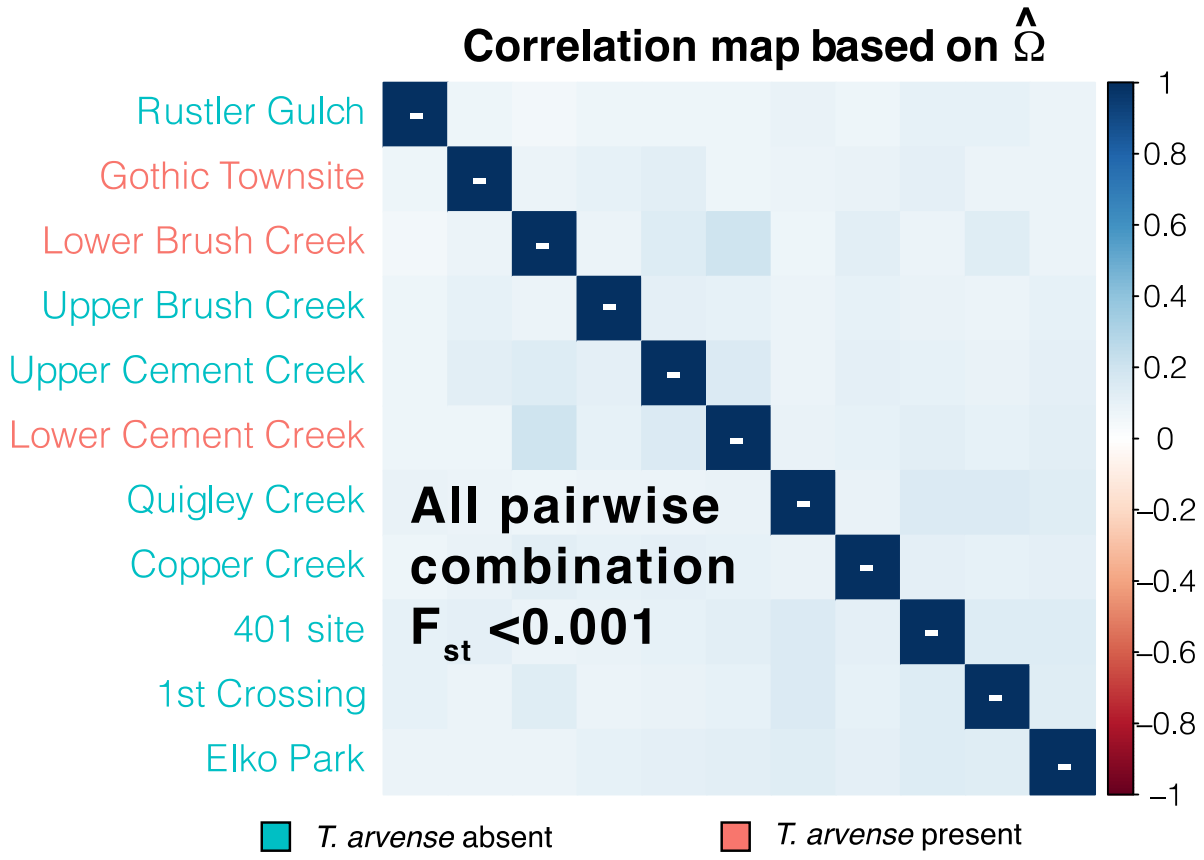

Figure A.2. Pairwise Fst between sites.

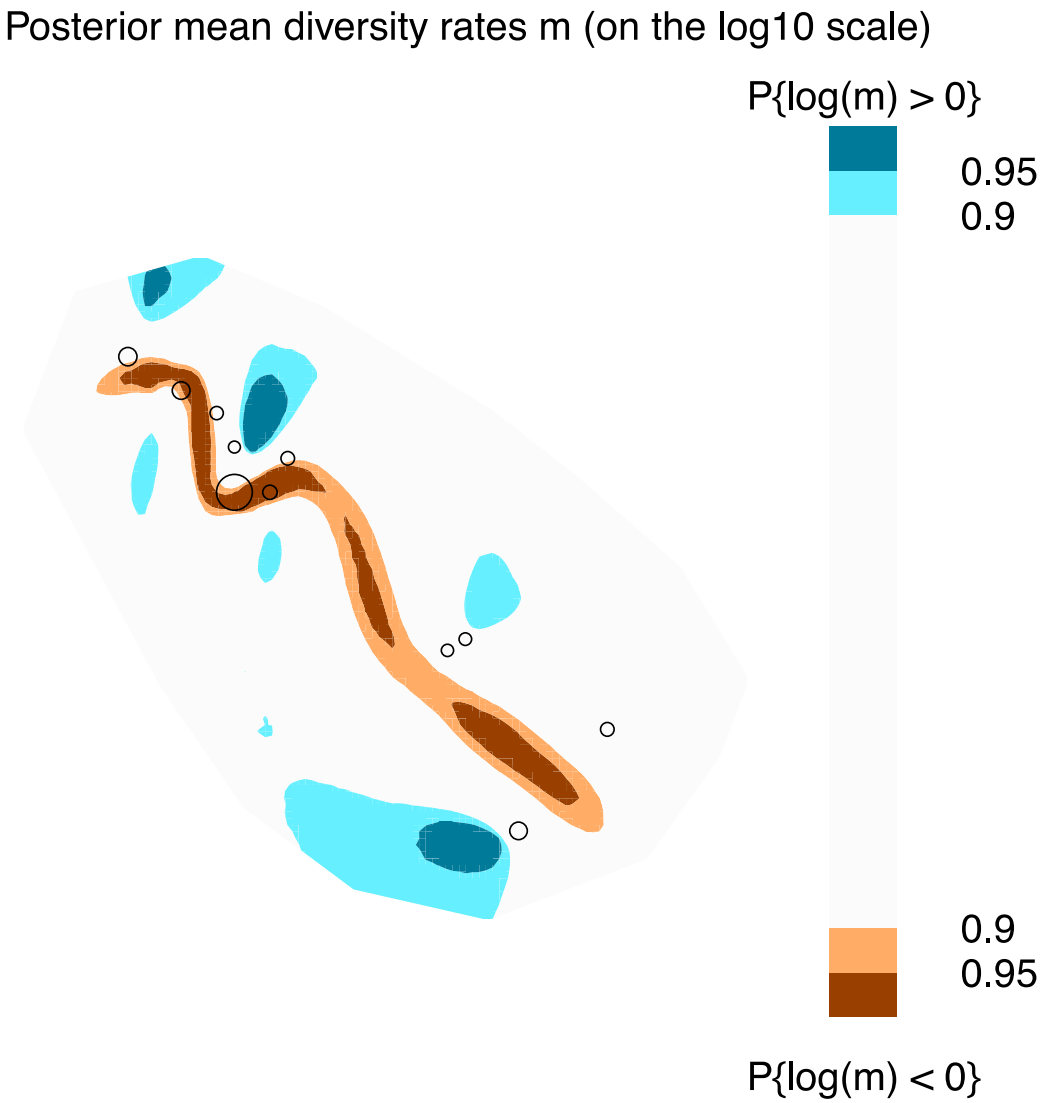

Figure A.3. Posterior migration rates.

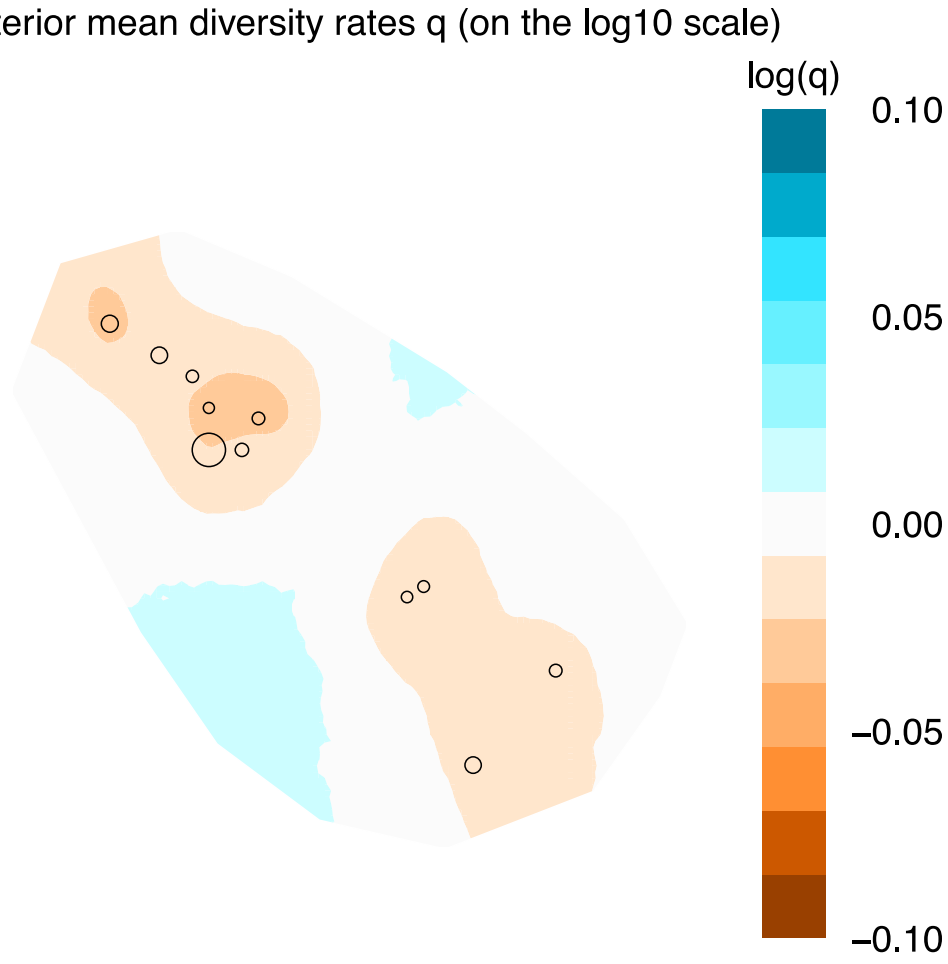

Figure A.4. Posterior diversity rates.

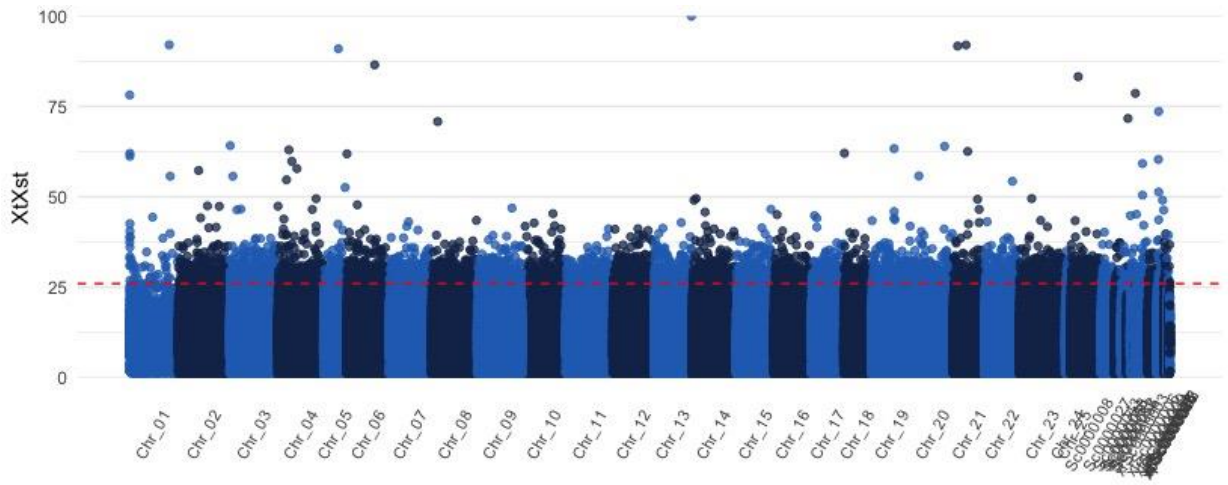

Figure A.5. Outlier SNPs associated with adaptive differentiation between areas with and without *T. arvense* at POD of XtX> 26.

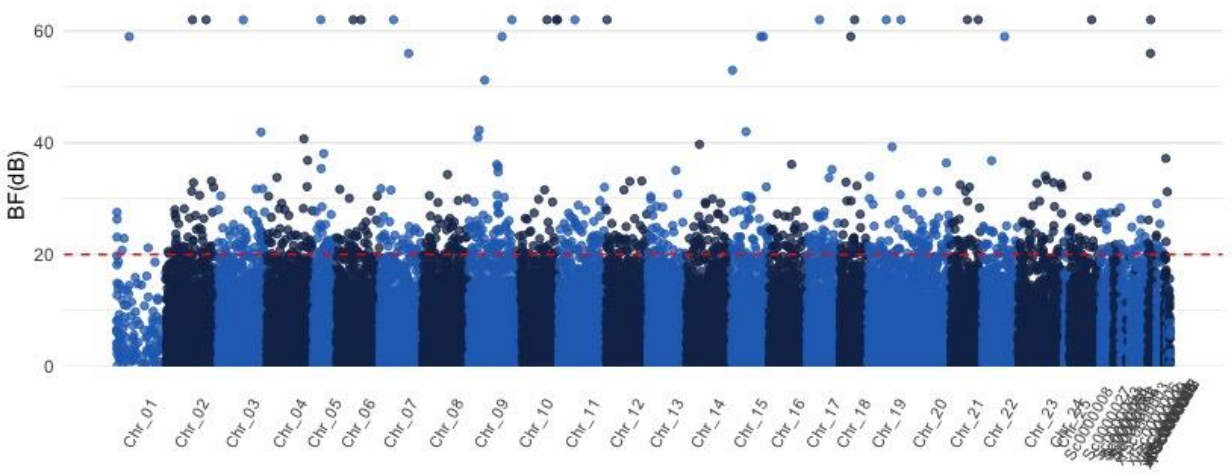

Figure A.6. Outlier SNPs associated with *T. arvense* at BF > 20.

APPENDIX B

CHAPTER 3 SUPPLEMENTARY INFORMATION

B.1 Tables

Table B.1. GO terms enriched in Cluster 1

TermID	Name
GO:0005975	carbohydrate metabolic process
GO:0006063	uronic acid metabolic process
GO:0006356	regulation of transcription by RNA polymerase I
GO:0006575	cellular modified amino acid metabolic process
GO:0006766	vitamin metabolic process
GO:0005996	monosaccharide metabolic process
GO:0006082	organic acid metabolic process
GO:0006811	monoatomic ion transport
GO:0055085	transmembrane transport
GO:0007205	protein kinase C-activating G protein-coupled receptor signaling pathway
GO:0007447	imaginal disc pattern formation
GO:0035222	wing disc pattern formation
GO:0007567	parturition
GO:0007588	excretion
GO:0008038	neuron recognition
GO:0008643	carbohydrate transport
GO:0009256	10-formyltetrahydrofolate metabolic process
GO:0009636	response to toxic substance
GO:0009804	coumarin metabolic process
GO:0009698	phenylpropanoid metabolic process
GO:0009812	flavonoid metabolic process
GO:0010578	regulation of adenylate cyclase activity involved in G protein-coupled receptor signaling pathway
GO:0007190	activation of adenylate cyclase activity
GO:0010677	negative regulation of cellular carbohydrate metabolic process
GO:0010817	regulation of hormone levels
GO:0014064	positive regulation of serotonin secretion

GO:0002027	regulation of heart rate
GO:0010700	negative regulation of norepinephrine secretion
GO:0014061	regulation of norepinephrine secretion
GO:0014062	regulation of serotonin secretion
GO:0032277	negative regulation of gonadotropin secretion
GO:0043300	regulation of leukocyte degranulation
GO:0043302	positive regulation of leukocyte degranulation
GO:0061178	regulation of insulin secretion involved in cellular response to glucose stimulus
GO:0061179	negative regulation of insulin secretion involved in cellular response to glucose stimulus
GO:1903524	positive regulation of blood circulation
GO:2000293	negative regulation of defecation
GO:0015711	organic anion transport
GO:0015851	nucleobase transport
GO:0015882	L-ascorbic acid transmembrane transport
GO:0034219	carbohydrate transmembrane transport
GO:0035461	vitamin transmembrane transport
GO:0070837	dehydroascorbic acid transport
GO:0019852	L-ascorbic acid metabolic process
GO:0021854	hypothalamus development
GO:0030324	lung development
GO:0030323	respiratory tube development
GO:0033008	positive regulation of mast cell activation involved in immune response
GO:0002888	positive regulation of myeloid leukocyte mediated immunity
GO:0033006	regulation of mast cell activation involved in immune response
GO:0035082	axoneme assembly
GO:0035482	gastric motility
GO:0035640	exploration behavior
GO:0035934	corticosterone secretion
GO:0001820	serotonin secretion
GO:0043303	mast cell degranulation
GO:0048243	norepinephrine secretion
GO:0050432	catecholamine secretion
GO:0051458	corticotropin secretion
GO:0140353	lipid export from cell
GO:0042560	pteridine-containing compound catabolic process
GO:0009396	folic acid-containing compound biosynthetic process
GO:0009397	folic acid-containing compound catabolic process
GO:0042559	pteridine-containing compound biosynthetic process

GO:0043949	regulation of cAMP-mediated signaling
GO:0043951	negative regulation of cAMP-mediated signaling
GO:0043950	positive regulation of cAMP-mediated signaling
GO:0044281	small molecule metabolic process
GO:0045922	negative regulation of fatty acid metabolic process
GO:0045943	positive regulation of transcription by RNA polymerase I
GO:0048521	negative regulation of behavior
GO:0048630	skeletal muscle tissue growth
GO:0050919	negative chemotaxis
GO:0051167	xylulose 5-phosphate metabolic process
GO:0042776	proton motive force-driven mitochondrial ATP synthesis
GO:1901159	xylulose 5-phosphate biosynthetic process
GO:0051179	localization
GO:0051180	vitamin transport
GO:0051867	general adaptation syndrome, behavioral process
GO:0048266	behavioral response to pain
GO:0051866	general adaptation syndrome
GO:0070482	response to oxygen levels
GO:0070633	transepithelial transport
GO:0071376	cellular response to corticotropin-releasing hormone stimulus
GO:0043435	response to corticotropin-releasing hormone
GO:0097305	response to alcohol
GO:0090069	regulation of ribosome biogenesis
GO:0090070	positive regulation of ribosome biogenesis
GO:1904223	regulation of glucuronosyltransferase activity
GO:2000232	regulation of rRNA processing
GO:2000234	positive regulation of rRNA processing

Table B.2. GO terms enriched in Cluster 2

TermID	Name
GO:0000003	reproduction
GO:0000909	sporocarp development involved in sexual reproduction
GO:0048608	reproductive structure development
GO:0048806	genitalia development
GO:0075259	spore-bearing structure development
GO:0001890	placenta development
GO:0001893	maternal placenta development
GO:0001892	embryonic placenta development
GO:0060135	maternal process involved in female pregnancy
GO:0060706	cell differentiation involved in embryonic placenta development
GO:0002237	response to molecule of bacterial origin
GO:0003170	heart valve development
GO:0003205	cardiac chamber development
GO:0003279	cardiac septum development
GO:0007507	heart development
GO:0060976	coronary vasculature development
GO:0006011	UDP-glucose metabolic process
GO:0009225	nucleotide-sugar metabolic process
GO:0022611	dormancy process
GO:0031665	negative regulation of lipopolysaccharide-mediated signaling pathway
GO:0002832	negative regulation of response to biotic stimulus
GO:0031664	regulation of lipopolysaccharide-mediated signaling pathway
GO:0032102	negative regulation of response to external stimulus
GO:0033078	extrathymic T cell differentiation
GO:0001779	natural killer cell differentiation
GO:0001865	NK T cell differentiation
GO:0030101	natural killer cell activation
GO:0042100	B cell proliferation
GO:0033082	regulation of extrathymic T cell differentiation
GO:0002683	negative regulation of immune system process
GO:0002695	negative regulation of leukocyte activation
GO:0002696	positive regulation of leukocyte activation
GO:0032814	regulation of natural killer cell activation
GO:0032823	regulation of natural killer cell differentiation
GO:0045577	regulation of B cell differentiation
GO:0045579	positive regulation of B cell differentiation

GO:0050865	regulation of cell activation
GO:0050866	negative regulation of cell activation
GO:0050867	positive regulation of cell activation
GO:0050869	negative regulation of B cell activation
GO:0051136	regulation of NK T cell differentiation
GO:1902107	positive regulation of leukocyte differentiation
GO:1903011	negative regulation of bone development
GO:1903708	positive regulation of hemopoiesis
GO:0035074	pupation
GO:0035152	regulation of tube architecture, open tracheal system
GO:0035209	pupal development
GO:0040011	locomotion
GO:0044706	multi-multicellular organism process
GO:0045682	regulation of epidermis development
GO:0045684	positive regulation of epidermis development
GO:0061050	regulation of cell growth involved in cardiac muscle cell development
GO:0061051	positive regulation of cell growth involved in cardiac muscle cell development
GO:0045834	positive regulation of lipid metabolic process
GO:0046889	positive regulation of lipid biosynthetic process
GO:0051674	localization of cell
GO:0060575	intestinal epithelial cell differentiation
GO:0060576	intestinal epithelial cell development
GO:0062013	positive regulation of small molecule metabolic process
GO:0071216	cellular response to biotic stimulus
GO:0071390	cellular response to ecdysone
GO:0006805	xenobiotic metabolic process
GO:0031663	lipopolysaccharide-mediated signaling pathway
GO:0035075	response to ecdysone
GO:0036314	response to sterol
GO:0036315	cellular response to sterol
GO:0071466	cellular response to xenobiotic stimulus
GO:0097306	cellular response to alcohol
GO:0090256	regulation of cell proliferation involved in imaginal disc-derived wing morphogenesis
GO:0008285	negative regulation of cell population proliferation
GO:0070663	regulation of leukocyte proliferation
GO:0070664	negative regulation of leukocyte proliferation
GO:0097628	distal tip cell migration
GO:1902884	positive regulation of response to oxidative stress

GO:1903355	negative regulation of distal tip cell migration
GO:0040013	negative regulation of locomotion
GO:1903354	regulation of distal tip cell migration
GO:2000146	negative regulation of cell motility
GO:1903624	regulation of DNA catabolic process
GO:1903625	negative regulation of DNA catabolic process
GO:1903627	regulation of dUTP diphosphatase activity
GO:0051340	regulation of ligase activity
GO:0051342	regulation of cyclic-nucleotide phosphodiesterase activity
GO:1903629	positive regulation of dUTP diphosphatase activity
GO:1904406	negative regulation of nitric oxide metabolic process
GO:1990654	sebum secreting cell proliferation
GO:0070661	leukocyte proliferation
GO:0090255	cell proliferation involved in imaginal disc-derived wing morphogenesis

Table B.3. GO terms enriched in cluster 4

TermID	Name
GO:0001906	cell killing
GO:0002376	immune system process
GO:0002682	regulation of immune system process
GO:0002684	positive regulation of immune system process
GO:0002697	regulation of immune effector process
GO:0002804	positive regulation of antifungal peptide production
GO:0032815	negative regulation of natural killer cell activation
GO:0032826	regulation of natural killer cell differentiation involved in immune response
GO:0032827	negative regulation of natural killer cell differentiation involved in immune response
GO:0050776	regulation of immune response
GO:0050778	positive regulation of immune response
GO:0050866	negative regulation of cell activation
GO:0002831	regulation of response to biotic stimulus
GO:0048584	positive regulation of response to stimulus
GO:0006952	defense response
GO:0008063	Toll signaling pathway
GO:0008592	regulation of Toll signaling pathway
GO:0009404	toxin metabolic process
GO:0009407	toxin catabolic process
GO:0009410	response to xenobiotic stimulus
GO:0009605	response to external stimulus
GO:0009607	response to biotic stimulus
GO:0009636	response to toxic substance
GO:0016998	cell wall macromolecule catabolic process
GO:0044036	cell wall macromolecule metabolic process
GO:0019227	neuronal action potential propagation
GO:0019730	antimicrobial humoral response
GO:0002440	production of molecular mediator of immune response
GO:0006955	immune response
GO:0006959	humoral immune response
GO:0009617	response to bacterium
GO:0031640	killing of cells of another organism
GO:0042742	defense response to bacterium
GO:0043207	response to external biotic stimulus
GO:0050830	defense response to Gram-positive bacterium

GO:0051701	biological process involved in interaction with host
GO:0098543	detection of other organism
GO:0019835	cytolysis
GO:0030913	paranodal junction assembly
GO:0031102	neuron projection regeneration
GO:0031341	regulation of cell killing
GO:0031343	positive regulation of cell killing
GO:0032689	negative regulation of type II interferon production
GO:0033631	cell-cell adhesion mediated by integrin
GO:0034113	heterotypic cell-cell adhesion
GO:0035640	exploration behavior
GO:0042181	ketone biosynthetic process
GO:0042268	regulation of cytolysis
GO:0042446	hormone biosynthetic process
GO:0044419	biological process involved in interspecies interaction between organisms
GO:0045161	neuronal ion channel clustering
GO:0045752	positive regulation of Toll signaling pathway
GO:0045919	positive regulation of cytolysis
GO:0060536	cartilage morphogenesis
GO:0003382	epithelial cell morphogenesis
GO:0031424	keratinization
GO:0070268	cornification
GO:0071205	protein localization to juxtaparanode region of axon
GO:0002175	protein localization to paranode region of axon
GO:0099612	protein localization to axon
GO:0071466	cellular response to xenobiotic stimulus
GO:0042178	xenobiotic catabolic process
GO:0071542	dopaminergic neuron differentiation
GO:0071554	cell wall organization or biogenesis
GO:1900150	regulation of defense response to fungus
GO:0002833	positive regulation of response to biotic stimulus
GO:0031349	positive regulation of defense response
GO:0032103	positive regulation of response to external stimulus
GO:1903859	regulation of dendrite extension
GO:1904386	response to L-phenylalanine derivative
GO:0097068	response to thyroxine
GO:0097069	cellular response to thyroxine stimulus
GO:1904387	cellular response to L-phenylalanine derivative

GO:1990089	response to nerve growth factor
GO:0097066	response to thyroid hormone
GO:0097067	cellular response to thyroid hormone stimulus
GO:1990090	cellular response to nerve growth factor stimulus

Table B.4. GO terms enriched in cluster 5

TermID	Name
GO:0001101	response to acid chemical
GO:0001816	cytokine production
GO:0001817	regulation of cytokine production
GO:0001819	positive regulation of cytokine production
GO:0006457	protein folding
GO:0006577	amino-acid betaine metabolic process
GO:0006578	amino-acid betaine biosynthetic process
GO:0006629	lipid metabolic process
GO:0006633	fatty acid biosynthetic process
GO:0001516	prostaglandin biosynthetic process
GO:0001676	long-chain fatty acid metabolic process
GO:0006082	organic acid metabolic process
GO:0006636	unsaturated fatty acid biosynthetic process
GO:0006644	phospholipid metabolic process
GO:0006690	icosanoid metabolic process
GO:0008610	lipid biosynthetic process
GO:0009395	phospholipid catabolic process
GO:0016042	lipid catabolic process
GO:0016053	organic acid biosynthetic process
GO:0019369	arachidonic acid metabolic process
GO:0033559	unsaturated fatty acid metabolic process
GO:0034638	phosphatidylcholine catabolic process
GO:0036151	phosphatidylcholine acyl-chain remodeling
GO:0036152	phosphatidylethanolamine acyl-chain remodeling
GO:0043651	linoleic acid metabolic process
GO:0044242	cellular lipid catabolic process
GO:0044255	cellular lipid metabolic process
GO:0044283	small molecule biosynthetic process
GO:0046337	phosphatidylethanolamine metabolic process
GO:0046338	phosphatidylethanolamine catabolic process
GO:0046394	carboxylic acid biosynthetic process
GO:0046456	icosanoid biosynthetic process
GO:0046457	prostanoid biosynthetic process
GO:0046470	phosphatidylcholine metabolic process
GO:0046475	glycerophospholipid catabolic process
GO:0046486	glycerolipid metabolic process

GO:0046503	glycerolipid catabolic process
GO:0072330	monocarboxylic acid biosynthetic process
GO:1901568	fatty acid derivative metabolic process
GO:1901570	fatty acid derivative biosynthetic process
GO:0007021	tubulin complex assembly
GO:0007205	protein kinase C-activating G protein-coupled receptor signaling pathway
GO:0007516	hemocyte development
GO:0007568	aging
GO:0008219	cell death
GO:0008340	determination of adult lifespan
GO:0009605	response to external stimulus
GO:0010259	multicellular organism aging
GO:0010876	lipid localization
GO:0015849	organic acid transport
GO:0019637	organophosphate metabolic process
GO:0046434	organophosphate catabolic process
GO:0019748	secondary metabolic process
GO:0032309	icosanoid secretion
GO:0071715	icosanoid transport
GO:1903963	arachidonate transport
GO:0033036	macromolecule localization
GO:0035233	germ cell repulsion
GO:0035234	ectopic germ cell programmed cell death
GO:0010623	programmed cell death involved in cell development
GO:0060561	apoptotic process involved in morphogenesis
GO:1902742	apoptotic process involved in development
GO:0038089	positive regulation of cell migration by vascular endothelial growth factor signaling pathway
GO:0042594	response to starvation
GO:0042810	pheromone metabolic process
GO:0043281	regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:0046717	acid secretion
GO:0051354	negative regulation of oxidoreductase activity
GO:0060326	cell chemotaxis
GO:0006935	chemotaxis
GO:0070542	response to fatty acid
GO:0033993	response to lipid
GO:0071105	response to interleukin-11

GO:0071348	cellular response to interleukin-11
GO:0097164	ammonium ion metabolic process
GO:0120254	olefinic compound metabolic process
GO:1901571	fatty acid derivative transport
GO:0006869	lipid transport
GO:0015908	fatty acid transport
GO:1901615	organic hydroxy compound metabolic process
GO:1902883	negative regulation of response to oxidative stress
GO:1905907	negative regulation of amyloid fibril formation
GO:1905906	regulation of amyloid fibril formation
GO:1990000	amyloid fibril formation
GO:2000027	regulation of animal organ morphogenesis
GO:2000378	negative regulation of reactive oxygen species metabolic process
GO:2001233	regulation of apoptotic signaling pathway
GO:0046668	regulation of retinal cell programmed cell death
GO:0046670	positive regulation of retinal cell programmed cell death