UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Building the bridge to herbivory: ecology, microbes, and adaptation

Permalink

https://escholarship.org/uc/item/31x8j0wd

Author

Aguilar, Jessica Maria

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Building the bridge to herbivory: ecology, microbes, and adaptation

by

Jessica Maria Aguilar

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

and the Designated Emphasis

in

Computational and Genomic Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Noah K. Whiteman, Chair

Professor Britt Koskella

Professor Craig Miller

Professor Michael Nachman

Spring 2024

Abstract

Building the bridge to herbivory: ecology, microbes, and adaptation

by

Jessica Maria Aguilar

Doctor of Philosophy in Integrative Biology

Designated Emphasis in Computational and Genomic Biology

University of California, Berkeley

Professor Noah K. Whiteman, Chair

Identifying the macro- and micro-evolutionary factors that create and maintain biodiversity is a central goal of evolutionary biology. It is clear that interactions between plant hosts and insect herbivores drive increased diversification rates in insects. Yet, there are also challenges to herbivory as a life history strategy. My dissertation research has focused on the evolution of herbivory and the adaptive leaps that must occur to survive in this trophic niche including behavioral, morphological, nutritional and the ability to detoxify a well-defended plant diet. This work is largely focused on adaptations in *Scaptomyza flava*, a representative of a lineage nested within the *Drosophila* that transitioned to herbivory ~10-15 million years ago.

First, I characterized the trophic niche of *S. flava*. We used natural history studies to determine if *S. flava* is a true herbivore or a cryptic microbe-feeder, given that the ancestral character state for the family Drosophilidae is likely microbe-feeding. We quantified oviposition substrate choice and larval viability across food types, including yeast media, decaying plants and fresh plants. We found that *S. flava* had a strong preference for fresh plants and did not lay eggs in alternate food options. Next, we described morphological traits related to feeding across putatively herbivorous and non-herbivorous drosophilids. We found an overall enlargement of the mouth hook of herbivores, with a reduction in the number of teeth along the hook. We also found no difference in gut length among *Scaptomyza* species studied. Next, we compared nitrogen isotope and sterol profiles to understand how nutrients were being integrated into the body. We found a reduction in the N_{15} isotope ratio in S. flava compared to non-herbivores, which is expected in a species with fewer trophic levels between itself and its food. Through these multiple lines of evidence, we confirmed that *S. flava* is an obligate herbivore of living mustard plants.

Next, I explored the relationship between *S. flava* and microbes that live in their gut. I tested the hypothesis that the gut microbiota of nascent herbivores can facilitate the initial switch to herbivory by providing detoxification services of plant defensive chemicals. This is called the gut microbial facilitation hypothesis. We first characterized

the gut microbiome of *S. flava*. We found that the life-stage of the fly influenced the microbial community in the gut the most, with species also playing a role. We then focused on identifying bacteria that are able to detoxify the common plant defensive chemical phenethyl isothiocynate (PEITC) using the SaxA enzyme. We isolated bacterial strains from the *S. flava* gut that are both resistant to and can detoxify PEITCs using SaxA. Lastly, we tested whether bacteria with *saxA* conferred benefits to non-herbivores when challenged with PEITC compared to bacteria without *saxA*. We found that the functional contribution of detoxification by one bacterial species was not enough to confer an advantage. However, future directions look toward testing the gut microbial facilitation hypothesis in this powerful tripartite system in other contexts and with a representative gut community.

Finally, I focused on testing the pre-adaptation hypothesis, which posits that adaptations of herbivorous flies that enable them to detoxify plant chemicals may also enable them to detoxify other xenobiotic chemicals, like pesticides. We have taken a comparative look at *cyp6g1* across related herbivorous and non-herbivorous drosophild species. *Cyp6g1* encodes a Cytochrome P450 monooxygenase that has been implicated in both insecticide resistance and phytochemical detoxification. We used a branch-site test to uncover signatures of positive selection along the branch leading to a subset of the *cyp6g1* copies in herbivores. We then focused on the seven functional *cyp6g1* copies that *S. flava* has in its genome. We also discovered a P-transposon inserted in the intergenic region between *cyp6g1b* and *cyp6g1c*. Transposable element insertions *cis* of detoxification genes tend to be related to an increase in the ability to detoxify. We determined whether this transposon was present in several pooled datasets and found that it was present in all three of the focal populations which spanned two sites across two years. This finding indicates that at the scale that we investigated, the P-transposon is being maintained in this genomic location. Increases in expression of genes *cis* of transposon insertions can be one way detoxification increases are obtained, so we measured differential expression of *cyp6g1* copies in response to plant defensive chemicals. We found an increase in expression of one copy (*cyp6g1h*), though not the copy directly *cis* of the P-transposon, indicating it is unlikely to have influenced this increase. Finally, we developed a homologous expression assay to measure the affinity of each Cyp6g1 copy across the clade to detoxify either indole glucosinolate derived plant defensive compounds or insecticides, with further results of this assay forthcoming.

Overall, I have worked to understand the major axes of adaptation enabling the evolution of herbivory (morphological, behavioral, nutritional and metabolism of plant toxins) with a main focus on detoxification either through endogenous adaptations or through interactions with gut bacteria.

For the two loves of my life, Cern and Quentin.

Acknowledgements

First, thanks to my advisor, Noah, who has created a space where I've been able to learn a ton about science and being a conscientious scientist. Grateful to him for taking a chance on me!

Thanks also to my committee members, Britt Koskella, Craig Miller and Micheal Nachman–I always learned a lot from their feedback.

My labmates have been the nucleus of support throughout my dissertation time. I would especially like to acknowledge fellow PhD students Nicolas Alexandre, Kirsten Verster, Julianne Pelaez, Hiro Suzuki, Rebecca Tarnopol, Diler Haji, and Tim O'Conner who I've learned so much from both scientifically and personally. Lab postdocs Rebecca Duncan, Teru Matsunaga, Cathy Rushworth, Marianna Karageorgi, Moe Bakhtiari, and Sam Chaturvedi, who have all shaped my journey in their own ways. Ben Goldman-Huertas, who I'm so happy to have worked with at the end of my tenure. And the one and only Susan Bernstein, who's role as "lab mom" has been a constant source of entertainment and scientific support.

My passion project throughout my time at UCB has been inclusiveIB and then iBio to follow. Jaemin Lee has been my steadfast partner in crime throughout, and I couldn't have dreamed of a better person to work with. Also Ana Lyons, Ali Bhatti, the whole iMCB crew, especially Lisa Eshun-Wilson. Monica and Carina, who deserve their own section (see next). And a huge thank you to the funding that has made the symposia possible throughout the years including Big C, Office for Graduate Diversity Dean's Graduate Diversity Innovation Fund, Gilliam Fellowship DEI allowance, Integrative Biology DEI Small Grants, and the Division of Biological Sciences.

Monica Albe and Carina Galicia have been huge supporters of inclusiveIB (and me!) from the start. I'm lucky to have been able to work so closely with both of you, and can't stress how much you both make all of our lives better as graduate students. Similarly, I'm grateful to Kate Chase, who has supported my efforts to complete a Designated Emphasis in Comp Bio.

Indispensable help from facilities around campus include the Oxford Tract Greenhouse, run with the utmost care by Tina Wistrom. The Cell Culture Facility has been a huge help with experiments for Ch 3, especially Alison Killilea, Willie S. Hercule, and Sara Sosa. The QB3 Mass Spectrometry Laboratory, especially Anthony Iavarone. And the folks who keep the VLSB and Weill hall running, especially Derek Apodaca and Igor Kligman.

I've been very lucky to have been supported by fellowships throughout my time as a PhD student, which made it possible to do so many "extras." I'd like to acknowledge the UC Berkeley Chancellor's Fellowship, NSF Graduate Research Fellowship, HHMI Gilliam Fellowship and my time on the Genetics Training Grant (especially Kathleen Ryan and Iswar Hariharan). Also, a Rosemary Grant award from the Society for the Study of Evolution and a summer research award from the Department of Integrative Biology supported some of the work described in Chapter 3.

And finally, I've been lucky enough to have been supported by my family and partner throughout this process. I can't thank them enough, and I definitely don't take the opportunity to do my PhD so close to home and family for granted. This has been especially important since my baby, Quentin, was born. The babysitters club has included my mother-in-law, May Saeteurn, my mom and sister, Lori and Sarah Aguilar, and Melanie Acuna. And a big thanks to Mike Aha and the Meridian Technical Services crew for allowing my husband, Cern, so much time to be home this winter. And thank you to Grandma and Grandpa Smith for sparking a love for plants and bugs in me early on, and MawMaw and PawPaw, who I know are rooting me on, just not from here.

CHAPTER 1 INTRODUCTION

In Chapter 1, I focused on understanding various traits that are necessary when evolving the ability to use plants as food and habitat. Plants provide sustenance for the most diverse lineage of Earth's named species, the insect herbivores, and they are also constantly evolving to protect themselves from damage. This instigates a tit-for-tat evolutionary pattern when herbivorous insects must evolve the ability to overcome those protections. Herbivores adapt through behavioral, morphological, and metabolic shifts in order to use plants as their hosts. In this chapter, I sought to understand whether the plant-associated fly, *Scaptomyza flava*, was using plants as a nutritional source by studying traits that underlie trophic shifts. The intended outcome of this study was to provide various lines of comparative evidence that, in addition to utilizing plants as habitat, which is easily observable, *S. flava* is also using plants as their main food source.

Insights into the evolution of herbivory from a leaf-mining fly

Jessica M. Aguilar[1]
Andrew D. Gloss[2*]
Hiromu C. Suzuki[1]
Kirsten I. Verster[1#]
Malvika Singhal[1]
Jordan Hoff[1]
Robert Grebenok[3]
Paul D. Nabity[2 %]
Spencer T. Behmer[4]
Noah K. Whiteman[1,5]

[1] Department of Integrative Biology, University of California-Berkeley, Berkeley, CA 94720, USA
[2] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
[3] Department of Biology, Canisius College, Buffalo, NY 14208, USA
[4] Department of Entomology, Texas A&M University, College Station, TX 77843, USA
[5] Department of Molecular & Cell Biology, University of California-Berkeley, Berkeley, CA 94720, USA

Corresponding Author: Noah K. Whiteman (whiteman@berkeley.edu)

Open Research Statement

All data and code underlying the statistical analysis in this paper can be found here: https://doi.org/10.5281/zenodo.7742074

Key Words/Phrases

Adaptation, *Arabidopsis*, *Drosophila*, herbivore, morphology, nutrition, oviposition, plant-insect interactions, *Scaptomyza,* sterols

Present addresses

[*] Department of Biology and Center for Genomics and Systems Biology, New York University, New York, NY 10003, USA

[#] Department of Biology, Stanford University, Palo Alto, CA 94301, USA

[%] Department of Botany and Plant Sciences, University of California-Riverside, Riverside, CA 92521, USA

Abstract

Herbivorous insects and their host plants comprise most known species on Earth. Illuminating how herbivory repeatedly evolved in insects from non-herbivorous lineages is critical to understanding how this biodiversity is created and maintained. We characterized the trophic niche of *Scaptomyza flava*, a representative of a lineage nested within the *Drosophila* that transitioned to herbivory ~10-15 million years ago. We used natural history studies to determine if *S. flava* is a true herbivore or a cryptic microbe-feeder, given that the ancestral character state for the family Drosophilidae is likely microbe-feeding. Specifically, we quantified oviposition substrate choice and larval viability across food-types, trophic-related morphological traits, and nitrogen isotope and sterol profiles across putatively herbivorous and non-herbivorous drosophilids. The results of these studies show that *S. flava* is an obligate herbivore of living plants. Paired with its genetic model host, *Arabidopsis thaliana*, *S. flava* is a novel and powerful system for exploring mechanisms underlying the evolution of herbivory, a complex trait that enabled the exceptional diversification of insects.

Introduction

Identifying the macro- and micro-evolutionary factors that create and maintain biodiversity is a central goal of evolutionary biology. Studies elucidating the ecological interactions between plants and their insect herbivores, the two most species-rich groups of macroscopic organisms on the planet, have been critical to addressing this challenge. It is well established that interactions between plant host and insect herbivores drive increased diversification rates in insects (Ehrlich & Raven, 1964; Mitter et al., 1988). Although 50% of insect species are herbivorous, only 30% of extant insect orders contain herbivores (Mitter et al., 1988; Schoonhoven et al., 2005). This paradox—that a minority of orders contain nearly half of all insect species—suggests that plants present challenging evolutionary hurdles for insects to initially overcome as food and habitat (Southwood, 1972).

There has been progress in understanding the genetic basis of herbivory, speciation, and host-plant specialization in herbivorous insects (Feder et al. 1988, McPheron et al. 1988, Gompert et al. 2012, Heliconius Genome Consortium 2012, Soria-Carrasco et al. 2014, Goldman-Huertas et al. 2015, Vertacnik and Linnen 2017). However, because herbivory evolved over 50 million years ago in the most diverse herbivorous insect lineages, the development of new model systems in which herbivory has evolved more recently is critical to understand broad patterns underlying the evolution of herbivory itself.

An ideal lineage in which to study the evolution of herbivory is one in which: (1) this trait has evolved relatively recently (so that close, non-herbivorous relatives can be used in comparative studies), (2) genomic tools are available (a well-annotated reference genome (Figure S1); amenability to germline transformation), (3) its generation time is relatively short (to facilitate laboratory and field studies), and (4) it

attacks a rapidly cycling host plant with its own suite of well-developed genetic tools (so that both sides of the plant-herbivore equation can be experimentally manipulated). The drosophilid fly *Scaptomyza flava* (Fallén, 1823) meets all of these criteria (Gloss et al., 2014; Goldman-Huertas et al., 2015; Matsunaga et al., 2022; Peláez et al., 2022; Whiteman et al., 2012, 2011).

The transition from microbe feeding to herbivory by the lineage containing *S. flava* is relatively recent (~10-15 million years before present; (Lapoint et al., 2013; Whiteman et al., 2012) among an array of insect species with sequenced genomes (Appendix 1: Figure S1). It is nested in a clade of well-studied, non-herbivorous drosophilid flies, including *Drosophila melanogaster*, and *S. flava* attacks the model plant *Arabidopsis thaliana* in nature and the laboratory (Mitchell-Olds, 2001; Whiteman et al., 2011). The adult females have leaf-puncturing ovipositors that they use to create feeding punctures and lay eggs in leaves, where larvae hatch and complete development as leaf miners (Peláez et al., 2022). The constitutive and jasmonate-induced defenses of *A. thaliana* provide some basal and inducible resistance against *S. flava* (Whiteman et al., 2012). However, the trophic niche of *S. flava*, a key aspect of its natural history, has not yet been empirically characterized.

S. flava and congeners are closely associated with living vegetative tissue of plants (Hackman 1959, Martin 2004), and can only be reared from living plants (Whiteman et al. 2011), but these factors alone do not eliminate the possibility that *S. flava* and its relatives are cryptic microbe feeders and thereby gain nutrition from this trophic level. The possibility that *S. flava* is gaining most of its nutrition from microbes on and within leaves is especially relevant since *Scaptomyza* is nested within an ancestral microbe feeding lineage in which most species consume microbes and associated decaying plant tissue. In addition, the gut microbial community in *S. flava* overlaps almost completely with the phyllosphere bacterial community of its host plant (O'Connor et al., 2014). This overlap in community composition is consistent both with an animal acquiring nutrients sufficient for development from these microbes and with an herbivore that may ingest microbes incidentally.

What makes an herbivore?

Herbivory is a complex organismal syndrome that involves a suite of inter-dependent adaptations. In the most fundamental sense, an herbivore is an organism that obtains most or all of its nutrients from living plant tissues. There are several subsidiary traits that make herbivory possible; these traits evolve from ancestral character states already shaped by acquiring nutrients from other sources.

For most immature herbivorous insects, initial access to food largely depends on maternal oviposition choices. These choices are informed by visual (Lyu et al. 2018, Nagaya et al. 2021) and chemosensory (Goldman-Huertas et al., 2015; Matsunaga et al., 2022) cues relaying information about environmental conditions. Although ovipositing females generally lay eggs where larval survival is expected to be high,

there are some instances where eggs are laid on less suitable hosts, for example, when more suitable hosts are rare, or the costs to egg laying are low (i.e. when a female has high fecundity) (Jaenike 1978). In these cases, oviposition behavior is one important factor that drives host shifts (Bernays and Chapman 1994).

Additionally, given the differences in the physical properties of food substrate and nutrient availability in plants, morphological adaptations to plant feeding are also common. Generally, herbivores tend to have longer gut lengths compared to non-herbivorous close relatives, which may increase digestion efficiency. This pattern has been observed in many different animal lineages, including crustations, insects, reptiles, fish, and mammals (Yang and Joern 1994, German and Horn 2006, Vervust et al. 2010, Griffen and Mosblack 2011, Karasov et al., 2011). However, in insects, gut length is variable. Generally, insect species that eat liquid food (i.e. rotten fruit, plant juices) have longer digestive tracts compared to species that eat solid food (i.e. intact leaves, animal tissue) (Richards and Davies 1977). Another commonly observed morphological adaptation that facilitates trophic niche shifts are changes in the mouthparts. In drosophilids, plastic changes in maxillary mouth hook dentation accompany shifts from microbe-eating to cannibalism (Vijendravarma et al. 2013), and in lizards, changes in dentition have also coincided with shifts from carnivory to herbivory (Vervust et al. 2010).

Although behavioral and morphological adaptations are associated with trophic niche shifts, herbivores survive primarily on nutrients supplied by their host plants. All organisms require a common set of essential nutrients, but microbes, plants, and animals differ in the ways these nutrients are assimilated into the body depending on their trophic niche. These differences in nutrient integration translate to differences in nutrient composition in organisms across trophic levels. Two salient nutrients in this regard are sterols (Behmer and David Nes 2003, Jing and Behmer 2020) and nitrogen (Fagan et al., 2002; Sterner & Elser, 2002). Sterols are lipids that are essential for insect development, but because insects cannot synthesize them *de novo*, they rely on food-sources to obtain them or supplement their diet (Clark & Block, 1959; Ikekawa et al., 1993; S. Li & Jing, 2020). Insect sterol profiles are shaped both by initial sterol profiles of food sources, as well as variation in the types of sterols that can be synthesized through dealkylation (Janson et al. 2009, Jing et al. 2012, 2013). Similarly, ratios of heavy to light nitrogen isotopes ($\delta^{15}N$) from protein and amino acids in food can be used as indicators of trophic level (Perkins et al. 2014). As fixed nitrogen moves through trophic levels, light isotopes (14N) are preferentially lost resulting in an enrichment of heavier isotopes (15N) in animals occupying higher trophic levels. When comparing insects that eat microbes to insects that eat plants, a relative enrichment in heavy nitrogen isotopes occurs (McCutchan et al. 2003).

To determine the feeding niche of *S. flava* (Hackman 1959), we compared a subset of trophic niche-related traits between *S. flava* and closely related microbe feeding relatives, including *S. pallida, S. hsui* and *D. melanogaster* (Figure 1)*. S. pallida* is in the subgenus *Parascaptomyza* shared a common acestor with *S. flava* ca. 24 to 27 mya *18-25 mya* and *S. hsui* is in the subgenus *Hemiscaptomyza* and shared a common acestor with *S. flava* ca. 24 to 27 mya (Lapoint et al., 2013; Matsunaga et al., 2022).

Neither of these subgenera of *Scaptomyza* are known to include herbivorous species, unlike the subgenus *Scaptomyza*, in which *S. flava* nested. The *Scaptomyza* subgenus also includes Caryophyllaceae feeder, *S. graminum* and additional Brassicaceae feeders, including *S. montana* and *S. nigrita* (Matsunaga et al., 2022). We typically rear both *S. pallida* and *S. hsui* in the laboratory from yeast-based fly media with rotting spinach added, suggesting that these two species retain the ancestral character state for the family and eat microbes found in decaying plant tissue. It is well understood that *D. melanogaster* mainly feeds on microbes, especially brewer's yeast (*Saccharomyces cerevisiae*) and relatives found in decaying fruit. Salient traits that we compared across this array of four drosophilid species included: 1) oviposition preference and ability to develop on alternative food types representing ancestral dietary substrates–yeast media and decaying plant vegetative tissue, 2) morphological measurements of trophic-related traits, and 3) the metabolic breakdown and integration of food products into the body (sterols and nitrogen isotope ratios). We hypothesized that if *S. flava* is herbivorous, these flies would show a strong oviposition preference for living plants as well as morphological (gut and mouthpart) and nutritive adaptations that differ from non-herbivorous close relatives.

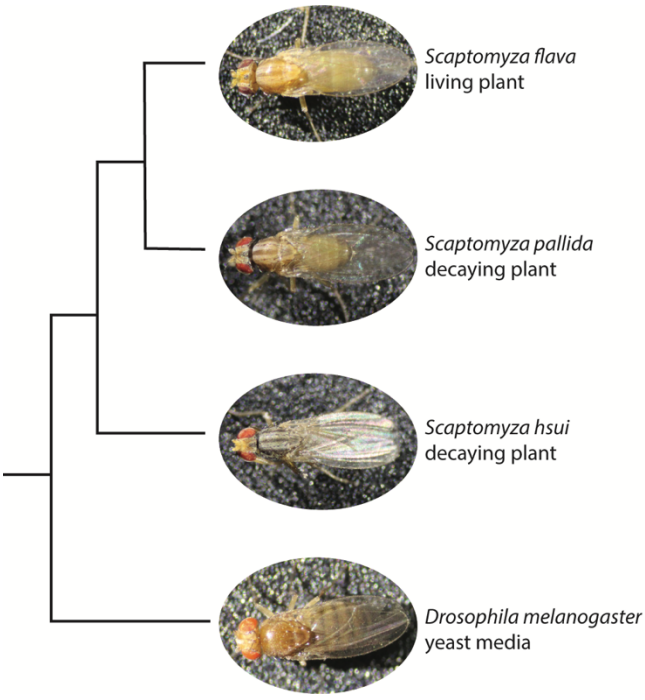


Figure 1: Phylogenetic relationship and laboratory food substrate *of S. flava, S. pallida*, *S. hsui*, and *D. melanogaster*. Tree topology from Peláez et al., 2022. Photos by Jessica M. Aguilar.

Materials and Methods

Oviposition choice and development assays

Oviposition preference assays: *S. flava* has exclusively been reared from living plants when collected in nature (Gloss et al., 2014; Whiteman et al., 2012, 2011). To determine if *S. flava* occupies a broader feeding niche than reported previously, we offered adult female flies a choice between living plants, decaying plants, and yeast-seeded media, and compared the choices of *S. flava* to that of microbe-feeding *D.*

melanogaster. For both living and decaying plant treatments, we grew *A. thaliana* (Col-0) plants to a six-week old rosette-stage as described in Whiteman et al. (2011) in Conviron growth chambers. To facilitate fungal decay that is often associated with *Drosophila* spp. colonization, we froze plants overnight to rupture cell walls, seeded the plant surface with baker's yeast (Fleischmann's) and subsequently placed them in a warmer environment (approx. 24°C) in the lab for three days before use (Chandler et al. 2012). We prepared standard *Drosophila* yeast media seeded with the same baker's yeast in petri dishes that approximated the diameter of the living and decaying plant rosette choices. Yeast media was made by the UC Berkeley fly food facility (Appendix 1: Supplementary methods). We randomly placed the three food substrates in a 30 x 30 cm mesh cage at 22°C and approximately 60% humidity. We then released four mated, 6–8-day old adult female flies from either an outbred *S. flava* colony collected from wild mustards (*Barbarea vulgaris and Turritis glabra*) near Dover, NH, USA or Canton-S *D. melanogaster* inside the cage. We counted eggs from each substrate after 24 hours. We compared differences in oviposition preference and viability among substrates using a permutation-based Kruskal-Wallis test in the R package *coin* (Hothorn et al., 2008; R Core Team, 2021). To ensure inferences were robust to the large number of tied values, we approximated significance against a null distribution constructed by Monte Carlo resampling (number of draws=$1e^7$) stratified by experimental block.

S. flava viability assays: To determine if *S. flava* requires living plants to complete development, we assayed larval viability on living plants, decaying plants, and yeast-seeded media, prepared and maintained as described above. Because we found that *S. flava* will only oviposit on living plants, we transferred 3–5-day old larvae from *A. thaliana* to each substrate by removing from fresh leaf material and placing on the surface of each foodtype. Larvae at this stage (approximately second instar) are able to burrow into new substrates independently. We tested nine replicates of six larvae for each substrate, and censused and removed adults that emerged at 13, 20, and 27 days. Differences among substrates were calculated as described above.

Comparative morphological measurements

We measured traits that are commonly associated with adaptations to herbivory including gut length, cephaloskeleton size, and mouth hook shape and tooth number in third instar *D. melanogaster, S. hsui, S. pallida,* and *S. flava* larvae. We reared flies on their normal lab diet: *D. melanogaster* on yeast media (UC Berkeley fly facility; Appendix 1: Supplementary methods), *S. hsui* on yeast media with decayed spinach (UC Berkeley fly food facility + Earthbound Farms brand spinach; Appendix 1: Supplementary methods), and *S. flava* on living *A. thaliana* plants. We performed all shape and size measurements in imageJ (Schneider et al. 2012) from photos taken with a scale bar and standardized magnification on a dissecting microscope. For all

comparisons, we ran ANOVA and Tukey's honestly significant difference test (Tukey's HSD) in R (R Core Team, 2021).

We measured wandering third instar larvae gut length and cephaloskeleton size measurements. We initially estimated this developmental stage by larval size and behavior. We then confirmed larval stage under a dissecting microscope by the presence of branched anterior spiracles (Martin 2012) and an orange hue on the posterior spiracles. We took body measurements when the larvae were in a fully extended state, which we initiated by submerging larvae in 70°C water for ~5 seconds (Godoy-Herrera et al. 1984). We measured midgut length, as defined from the proventriculus to the Malpighian tubules, by laying the organ such that it was in a single plane, aided by a glass coverslip, while avoiding stretching or compressing. We isolated cephaloskeletons from soft tissue using jeweler's forceps and took measurements from the tip of the oral process of the mouth hook to the bridge of the posterior process. We divided gut length and cephaloskeleton length by body length to account for variation caused by larvae size.

For mouth hook tooth quantification and shape measurements, we collected the anterior-most section of the third instar mouth hook from the puparium. We first soaked specimens in 70% ethanol for a minimum of 5 minutes to aid in dissection. We mounted right and left sides of each mouth hook separately on glass slides with Fisher Scientific Permount mounting media. Counting and measurements were done on the side of the specimen that was more clearly prepared and/or photographed, or if both sides were of equal quality, one was randomly chosen. We used landmarks as defined in Wipfler et al. (2013). We took mouth hook height as the distance from the ventral process to the dorsal process, and length was defined as the distance between the ventral process to the oral process. We calculated the ratio of height to length to get a sense of the shape distribution of the mouth hook (i.e., whether they were long and thin versus more robust in width compared to height). When counting tooth number, we identified the details of the teeth using a combination of photographs and visual confirmation under 40x magnification. Tooth size was calculated by subtracting the distance between the base of the longest tooth to the top of the mouth hook from the distance between the tip of the tooth to the top of the mouth hook. To depict shifts in morphology across species, we traced representative mouth hook images in Adobe Illustrator.

Sterol profiling

To better understand differences in sterol utilization from food sources, we extracted and characterized the sterol types and quantities present in *D. melanogaster, S. hsui,* and *S. flava* and the food on which they are reared in the lab. We chose these three species to represent the range of food types of *Scaptomyza* and close relatives. First, we reared flies on their normal lab diet: *D. melanogaster* on yeast media (Bloomington; Appendix 1: Supplementary methods), *S. hsui* on yeast media with decayed spinach (UC Berkeley fly food facility + Earthbound Farms brand spinach; Appendix 1: Supplementary methods), plain decayed spinach for comparison and *S. flava* on living

A. thaliana leaves. We added previously frozen, thawed store-bought spinach to the top of a subset of the yeast media vials. We left all food preparations at room temperature for 7 days, during which time the yeast media with spinach and the plain spinach were able to decay. *A. thaliana* plants were approximately 30 days old, and were grown as above.

We weighed and homogenized 21 individual *S. hsui* larvae, 13 individual *S. flava* larvae, a pool of 90 *D. melanogaster* puparia (Martin 2015) and samples of prepared food that had not been eaten. We processed these samples as per Jing et al., 2013 with slight modification. We extracted and processed free and base hydrolysable sterols, and glycosylated sterols (Appendix 1: Supplementary methods, (Jing et al. 2013, Feng et al. 2015). We conjugated free sterols and sterols freed following saponification and acid hydrolysis using a BSTFA+TMCS, 99:1 (Sylon BFT) Kit (Supelco, Bellefonte PA). We re-suspended samples in 800 µl of conjugation solution and hexane (4:1 v/v) and incubated for 12 h in the dark at room temperature. Following incubation, we added 2 ml of hexane and 2 ml of 70% methanol:water (v/v) and mixed the sample by pipetting. We removed the hexane fraction, evaporated to dryness under nitrogen. We subjected all fractions to GC-MS analysis.

We plotted the average sterol composition for *S. hsui* and *S. flava*, and the pooled composition for *D. melanogaster* in R (R Core Team, 2021). We calculated the Bray-Curtis dissimilarity between all samples using the *vegan* package (Oksanen et al. 2022). This measure takes both the presence of sterol compounds and the amounts of sterol compounds into account.

Isotope profiling

To measure $\delta^{15}N$ in *S. flava,* we reared flies in replicate cages with ample *A. thaliana* Col-0 rosettes grown to approximately 6 weeks as described above. We collected plant tissue immediately prior to exposure to adult *S. flava*, and again after 10 days of feeding using leaf tissue adjacent to any mines. We collected 25 third instar larvae for each of three replicates at 10 days and at 21 days. Immediately after collection, we dried tissues at 50°C, ground to a fine powder, and combusted using an Elemental Combustion System (Costech Analytical Technologies, Valencia, CA) coupled to a continuous-flow gas-ratio mass spectrometer (Thermo Finnigan Delta PlusXL, San Diego, CA). We quantified $\delta^{15}N$ using acetanilide standard for elemental concentration and IAEA-N-1 and IAEA-N-2. Precision is better than ± 0.2 for $\delta^{15}N$ (1 s.d.), based on repeated internal standards.

To compare *S. flava* $\delta^{15}N$ trophic enrichment with other herbivorous and non-herbivorous organisms, we performed a simple search in PubMed (in 2013) of “herbivore” “insect”, “trophic”, and “isotope” but only those publications with raw or summary data available on $\delta^{15}N$ for both consumer and food/host were used. We found additional sources by screening papers that cited those found in the original search. We restricted our search to fluid-feeding and defoliating herbivores. For both literature and

lab derived measurements, we calculated $\delta^{15}N$ trophic enrichment by subtracting diet $\delta^{15}N$ from insect $\delta^{15}N$. We compared $\delta^{15}N$ trophic enrichment values of herbivores and non-herbivores found in the literature using an ANOVA run in R (R Core Team, 2021)).

Results

S. flava oviposition choice and development assay

In a three-way choice test offering equal proportions of approximately 6-week-old, living *A. thaliana* rosettes, decaying *A. thaliana* rosettes, and yeast-seeded nutrient media, *S. flava* laid eggs exclusively in living leaves (N = 34 total eggs across 10 out of 17 trials, chi-squared = 19.767, p = $4.82e^{-5}$, Kruskal-Wallis test) and *D. melanogaster* exclusively in yeast media (Figure 2a ; N = 130 total eggs across 11 out of 30 trials, chi-squared = 21.902, p = $1.87e^{-5}$, Kruskal-Wallis test).

Next, we transferred second instar *S. flava* larvae onto each of the same three substrates. Larvae only completed development to adults in living plant tissue (Figure 2b; N = 9 replicates per substrate, with 6 larvae per replicate, chi-squared = 24.712, P = $9e^{-7}$, Kruskal-Wallis test).

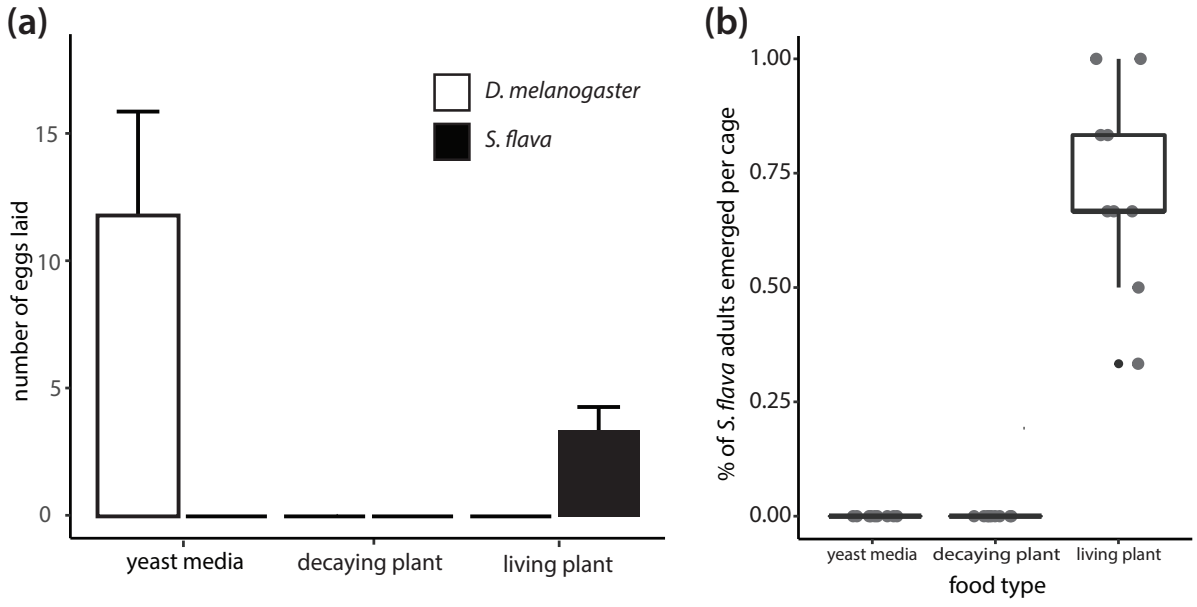


Figure 2: Choice and viability assays confirm that *S. flava* oviposition and development occur on living plants **(a)** Oviposition choice trials. Eggs laid by adult *S. flava* and *D. melanogaster* females in each of three substrates were counted after 24 hours. **(b)** *S. flava* development assay. Per replicate, six second instar *S. flava* were transplanted from living plants to individual living plants, decaying plants or yeast media tubes. The midline of the box plot represents the median, upper box line represents the third quartile and lower line represents the first quartile. The whisker terminates at either the most extreme value or one and a half times the inter quartile range with the most extreme value represented as a dot.

Comparative morphological measurements

We compared gut length divided by body size for four species (*D. melanogaster, S. hsui, S. pallida, S. flava*). We rejected the null hypothesis that average gut length was the same across species (Figure 3a; Appendix 1: Tables S2 and S3; P = <0.001). Further, after correcting for multiple comparisons with a Tukey's HSD test, we found *D. melanogaster* had the largest gut/body ratio compared to the three other species. All ANOVA and Tukey's HSD results can be found in Appendix 1: Tables S2 and S3. Next, we compared cephaloskeleton length divided by body size (Figure 3b). We found no differences between *D. melanogaster* and *S. hsui* or *S. pallida* and *S. flava*, but *S. pallida* and *S. flava* had larger cephaloskeletons than *D. melanogaster* and *S. hsui* (Appendix 1: Tables S2 and S3; P = <0.001). Figures 3d-f show comparisons of the anterior most section of the mouth hook. First, we compared the ratio of mouth hook height to length (Figure 3d). We found no differences between *D. melanogaster* and *S. pallida* or *S. hsui* and *S. flava*, but *S. hsui* and *S. flava* had mouth hooks with a longer length from top to bottom compared to *D. melanogaster* and *S. pallida* (Appendix 1: Tables S2 and S3; P = <0.001).

Finally, we quantified the tooth number and tooth size. Generally, tooth size had an inverse relationship to tooth number, where species with many teeth also had smaller teeth and species with fewer teeth had larger teeth. On average, *D. melanogaster* had the largest number of teeth, followed by *S. pallida*, *S. hsui*, and *S. flava* (Figure 3e; Appendix 1: Tables S2 and S3; P = <0.001). Across the four species measures, *S. pallida* and *S. hsui* had similarly sized teeth that fell between *D. melanogaster*'s smaller and *S. flava's* larger teeth (Figure 3f; Appendix 1: Tables S2 and S3; P <0.001). Generally, *S. flava* had fewer, larger teeth, *D. melanogaster* had many small teeth, and *S. hsui* and *S. pallida* fell in the intermediate space for both traits.

Sterol Profiling

We obtained values for the composition of sterols in laboratory food sources (Figure 4a) and flies that represent the feeding niches of *S. flava* and its close, non-herbivorous relatives, *D. melanogaster* and *S. hsui* to determine whether there were differences in sterol processing between these species (Figure 4b). To better understand differences between these profiles, we calculated the Bray-Curtis dissimilarity (BCD) of both the food and fly sterol profiles (Figure 4c). Decaying spinach was the most dissimilar from the other three food types, with scores above 0.9, and unsurprisingly, decaying spinach + yeast and plain yeast media were the most similar (Figure 4a, 4c; BCD = 0.26). For fly sterol profiles, *S. hsui* and *D. melanogaster* were the most similar (BCD = 0.6), and *S. flava* and *D. melanogaster* were the least similar (Figure 4b, 4c; BCD = 0.76). It can be

inferred that sitosterol dealkylating to cholesterol was most common in flies eating plant material (Figure 4d).

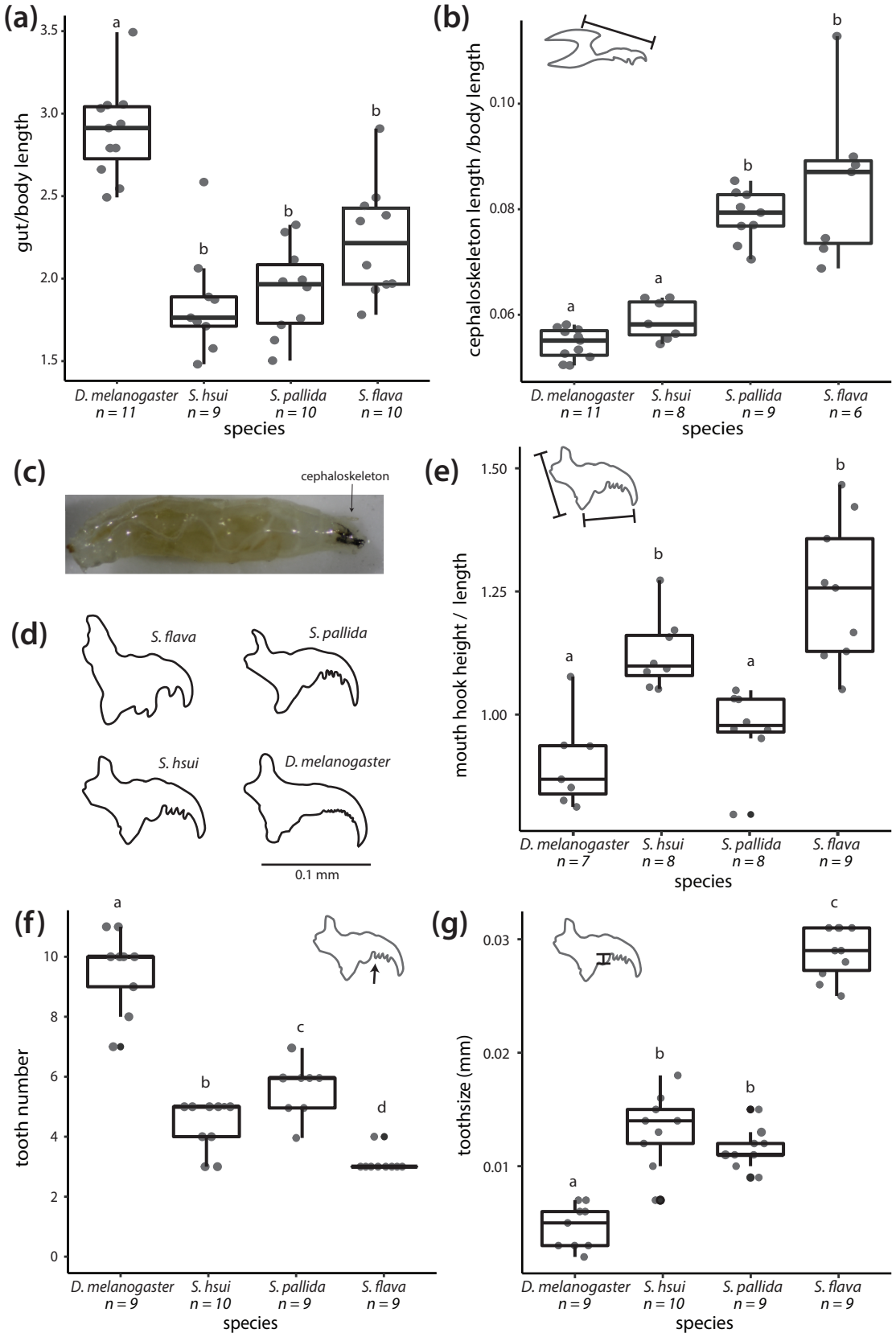

(a)
gut/body length
3.5
3.0
2.5
2.0
1.5
a
b
b
b
D. melanogaster n = 11
S. hsui n = 9
S. pallida n = 10
S. flava n = 10
species
(b)
cephaloskeleton length /body length
0.10
0.08
0.06
a
a
b
b
D. melanogaster n = 11
S. hsui n = 8
S. pallida n = 9
S. flava n = 6
species
(c)
cephaloskeleton
(d)
S. flava
S. pallida
S. hsui
D. melanogaster
0.1 mm
(e)
mouth hook height / length
1.50
1.25
1.00
a
b
a
b
D. melanogaster n = 7
S. hsui n = 8
S. pallida n = 8
S. flava n = 9
species
(f)
tooth number
10
8
6
4
2
0
a
b
c
d
D. melanogaster n = 9
S. hsui n = 10
S. pallida n = 9
S. flava n = 9
species
(g)
toothsize (mm)
0.03
0.02
0.01
a
b
b
c
D. melanogaster n = 9
S. hsui n = 10
S. pallida n = 9
S. flava n = 9
species

Figure 3: Morphological measurements of characters often associated with herbivory. Data points are offset on the x-axis to aid in visualization. Letters above bars indicate significant differences. For boxplots in panels a, b, d, e and f, the midline of the box plots represent the median, upper box line represents the third quartile and lower line represents the first quartile. The whisker terminates at either the most extreme value or one and a half times the inter quartile range with the most extreme value represented as a dot. **(a)** Length of gut relative to body length. **(b)** Length of cephaloskeleton relative to body length; inset photo of a *Drosophila melanogaster* larvae with cephaloskeleton labeled. **(c)** Representative depiction of mouth hook shape, mean tooth number and tooth size and shape. Photo by Jessica M. Aguilar. **(d)** Ratio of mouth hook height (here defined as the ventral process of the mouth hook (VPMH) to the dorsal process of the mouth hook) to the length (here defined as the VPMH to the oral process of the mouth hook). **(e)** Mouth hook tooth number. **(f)** Length of the longest tooth.

Isotope profiling

We profiled nitrogen isotope composition in *S. flava* and its *A. thaliana* Col-0 host plants and compared these to additional $\delta^{15}N$ profiles from the literature to determine whether they were more like an herbivore or non-herbivore $\delta^{15}N$ trophic enrichment (insect $\delta^{15}N$ – diet $\delta^{15}N$). Insects feeding on rotting, microbe-rich substrates typically exhibited a two-fold greater trophic enrichment of $\delta^{15}N$, than herbivores (Figure 4e; df = 1, F = 5.51, P = 0.023). Similarly, $\delta^{15}N$, was depleted in *S. flava* relative to its host plants, and $\delta^{15}N$ trophic enrichment was much lower than that of insects feeding on microbe-rich substrates.

Discussion

The evolution of herbivory is a key innovation that has spurred biodiversity in many animal lineages, but especially the most diverse animal group, plant feeding insects (Ehrlich & Raven, 1964; Mitter et al., 1988). Understanding adaptations involved in overcoming the physical and nutritional barriers that plants provide can help us better understand the evolution of this trophic niche. Building on previous work, in this study, we use a natural history approach to investigate the feeding habit of *S. flava*, an herbivorous, drosophilid fly that recently evolved from closely related non-herbivorous species. We quantified several key characteristics that might aid *S. flava* in plant feeding, including behavior, morphology, and nutritional integration. We compared these traits to three closely related, non-herbivorous drosophilid species.

The genus *Scaptomyza* is nested within the paraphyletic *Drosophila* clade. The species *S. flava* attacks the model plant *A. thaliana*, and well annotated reference genomes are available for several *Scaptomyza* congeners (B. Y. Kim et al., 2021; Peláez et al., 2023). Although an antagonistic ecological relationship between *Scaptomyza* species and its host plants has been characterized (Hackman, 1959; N. A. Martin, 2004; Whiteman et al., 2011) the question of its nutritional niche has not been addressed previously. Here, by comparing traits commonly associated with an herbivorous lifestyle, we found that *S. flava* is an obligate herbivore that consumes and acquires nutrients from living plant tissue.

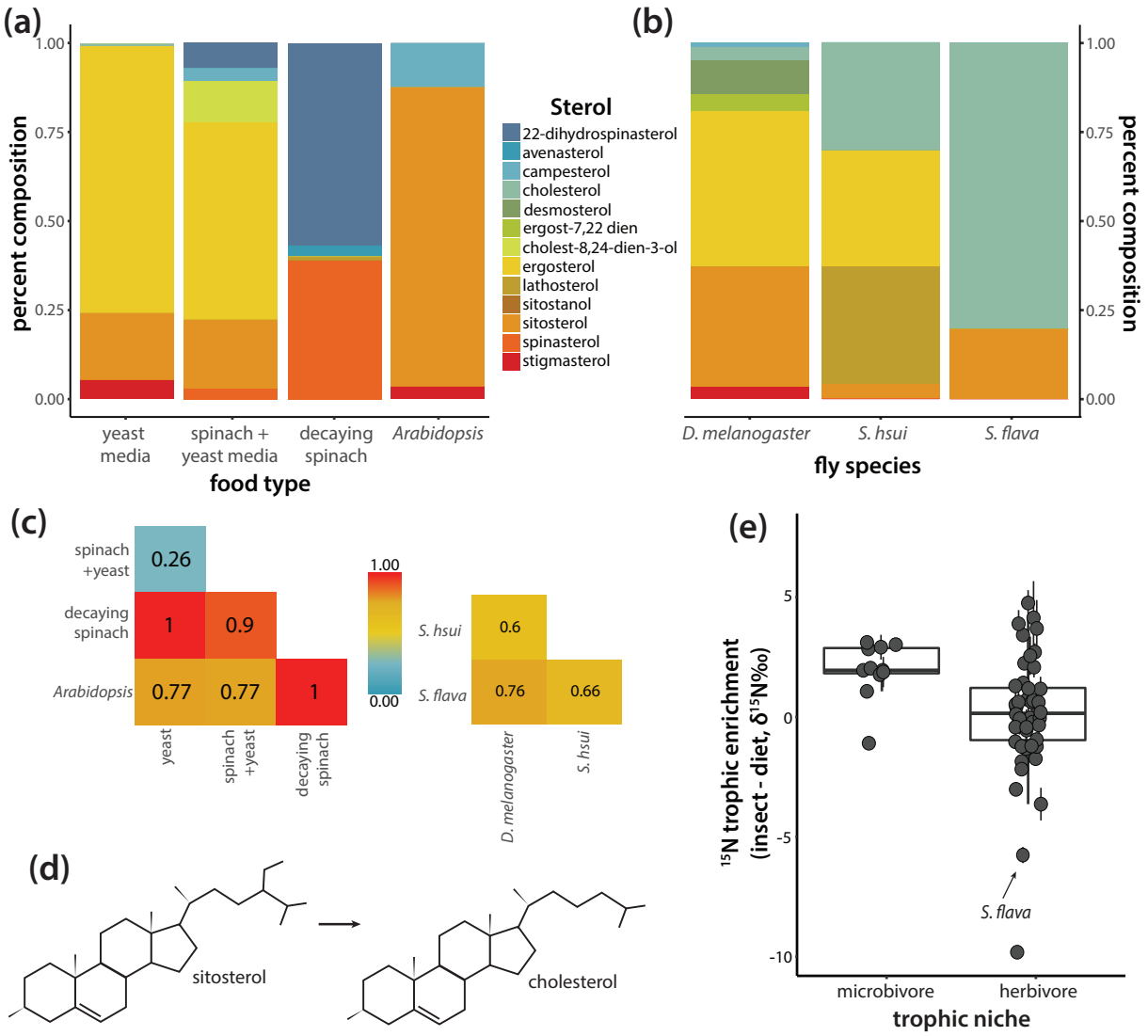


Figure 4: Sterol and nitrogen isotope profiles show *S. flava* integrates nutrients from plant leaves **(a)** Sterol composition of laboratory-reared fly food **(b)** Sterol composition of flies reared in the laboratory. *D. melanogaster* eats yeast media, *S. hsui* eats yeast media with decaying spinach, *S. flava* eats living *A. thaliana* plants. **(c)** Bray-Curtis dissimilarity measure of sterol composition of fly food and flies. A 1 is the most dissimilar composition, 0 is completely alike. **(d)** Dealkylation of sitosterol to cholesterol. Sitosterol is the most common type of sterol in plants. **(e)** Trophic enrichment of heavy nitrogen isotopes ($\delta^{15}N$, the ratio of ^{15}N:^{14}N) in insects relative to their feeding substrate, for insects consuming living plant material vs. decaying, microbe-rich plant material. Each circle indicates an averaged observation for a unique combination of insect species and host species or substrate +/- SEM. Data points are offset on the x-axis to aid in visualization. The midline of the box plot represents the median, upper box line represents the third quartile and lower line represents the first quartile. The whisker terminates at either the most extreme value or one and a half times the inter quartile range with the most extreme value represented as a dot. All data except *S. flava* measurements are from the literature. N insects eating decaying plants = 11, N insects eating living plants (excluding *S. flava*) = 44; P = 0.023.

Since early instar *S. flava* larvae are relatively confined to the location in which they were oviposited, oviposition behavior is key to understanding trophic niche shifts (Bernays and Chapman 1994). In our choice and development trials, *S. flava* showed a strong preference for living plants over decaying plants or yeast media when ovipositing; eggs were laid only on living plants (Figure 2a). This is consistent with the observation that adult *S. flava* antennae are responsive to volatiles from leaves but far less so from yeast in olfactory and behavioral assays (Goldman-Huertas et al., 2015; Matsunaga et al., 2022). After transplanting larvae to the three food types (living plants, decaying plants and yeast media), we found a similar trend for larval development to adulthood-adult flies only emerged from living plants (Figure 2b). In contrast, *D. melanogaster* showed a strong preference for yeast media, with eggs laid only on that food resource.

Morphological adaptations to different food types in multiple life history stages are also associated with trophic shifts. There have been studies exploring morphological adaptations in the ovipositors of *S. flava* adults that oviposit in leaves (Peláez et al., 2022), and here, we explored morphological adaptations in the herbivorous larval stage of *S. flava*. We used closely related, non-herbivorous groups (*S. pallida, S. hsui,* and *D. melanogaster*) as comparisons. Contrary to the commonly found pattern of longer gut lengths in herbivores compared to non-herbivores, gut length to body ratio was longer in non-herbivorous *D. melanogaster* compared to the other species studied, including the herbivorous *S. flava* (Figure 3a). This is likely related to the fact that they feed extra-orally and effectively drink their food; insects that drink their food tend to have longer alimentary canals (Richards and Davies 1977). Cephaloskeleton length was not correlated with herbivory, but we did find morphological differences in mouth hook shape and tooth size between herbivorous and non-herbivorous flies. Herbivorous *S. flava* had fewer, larger teeth on its mouth hooks, which are the mouth organs that mechanically gather food during eating (Figures 3b-f). The decrease in tooth number accompanying an increase in tooth size in herbivores is consistent with their utilization of a tougher food source and parallels what was found in carnivorous *D. melanogaster*, which showed a plastic reduction in mouth hook tooth number during shifts from yeast-substrate feeding to carnivory (Vijendravarma et al. 2013). However, even if putatively herbivorous species show strong oviposition preferences for living plants or morphological shifts consistent with herbivory, it is still possible that these species are obtaining most of their nutrients from microbes living in the leaf. One example of this is *D. sechellia,* which feeds on fresh fruits, but also relies on microbes for some nutritional components (Bing et al., 2018). To further understand the source of nutrition for these flies, we measured two metabolic outcomes: sterol profiles and nitrogen stable isotope profiles.

How insects metabolize ingested dietary sterols can provide unique insights to host plant use. For example, the profile of sterols in the bodies of insects is linked to differences in sterol profiles found in insect food, as well as adaptations in dealkylation ability of the species (Janson et al. 2009, Thompson et al. 2013). By comparing the sterol profiles in food to sterol profiles in insects using Bray-Curtis dissimilarity calculations, we obtained a combined picture of these two contributing factors (Figures 4a-c). When comparing herbivores to non-herbivores, we found differences in the

efficiency of dealkylation of sitosterol from food sources to cholesterol in insect bodies. Herbivorous *S. flava* had the most dissimilar sterol profile compared to its close non-herbivorous relatives (Figure 4c). In general, *S. flava* was the most efficient at dealkylation of sitosterol, but *S. hsui* is also able to dealkylate this common plant sterol (Figure 4a-d). However, *S. hsui* eats decaying leaves and in the lab is fed a mixture of decaying spinach and yeast media. The laboratory-fed diet of the two microbe feeding flies, *S. hsui* and *D. melanogaster,* contain the most similar sterol profiles, although the sterol profiles of these two insects are still quite different (Figure 4c). Unsurprisingly, *D. melanogaster* and *S. hsui* both contained ergosterol, which is ubiquitous in fungi; *S. hsui* can dealkylate the spinach-exclusive sterol spinasterol to lathosterol. Future studies into the mechanistic and evolutionary underpinnings of this are necessary to further interpret the significance of these results as well as to rule out the selective use of cholesterol by these herbivorous larvae (Behmer 2017, Jing and Behmer 2020).

Comparisons of stable isotope composition of nitrogen between insects and putative dietary substrates offer insight into nutrient acquisition: lighter isotopes are preferentially lost through excretion and metabolic activity relative to heavier ones as nutrients move up trophic levels, leading to an enrichment of heavy relative to light nitrogen isotopes ($\delta^{15}N : \delta^{14}N$) with each trophic step (Perkins et al. 2014). We found $\delta^{15}N$ enrichment in insect bodies that reflected the additional trophic step between insect and dietary substrate in microbe feeding insects (i.e., rotting plant tissue → microbe → insect, vs. living plant → insect) (Spence and Rosenheim 2005). Although there is overlap in $\delta^{15}N$ enrichment signatures between plant and microbe feeders, increased depletion of $\delta^{15}N$ is typical for phloem-feeding herbivorous insects (Wilson et al. 2011, Perkins et al. 2014), and is inconsistent with a diet reliant on microbes. Our findings suggest that *S. flava* derives nutrients from living plant tissue (Figure 4e). Although we have found evidence that nutrients from plants (and not microbes) are primarily assimilated into flies during development, there is still the possibility that microbes in the gut may play other roles in fly fitness, like through interactions with plant defenses. This is further supported by previous findings that larval feeding and growth rates are not directly enhanced by elevated bacterial loads in their diet, although larvae benefit indirectly from bacterial suppression of plant defenses (Groen et al. 2016).

Overall, the results of our experiments and comparative studies framed by natural history show that *S. flava* is an obligate herbivore which utilizes nutrients from living plants to survive. The *S. flava-A. thaliana* interaction system has the potential to facilitate our understanding of how most of the macroscopic biodiversity has evolved and is maintained, as well as fundamental mechanistic questions at the plant-herbivore interface. Future studies are needed to disentangle the adaptive leaps required at the transition to herbivory, as well as the ways herbivorous species diversity is created and maintained once the hurdles to herbivory have been overcome.

Acknowledgments

The authors would like to thank J.N. Pelaez, R. P. Duncan and M. Karageorgi for assistance with insect care, N.M. Alexandre, C. Miller, and M. Nachman for discussions on morphological measurements and manuscript content, and two anonymous reviewers for constructive feedback on an earlier version of this manuscript. This work was supported by a grant from the NIGMS (R35GM119816) to N.K.W, HHMI Gilliam Fellowship (GT13540) and National Science Foundation Graduate Research Fellowship (DGE 1752814/DGE 2146752) to J.M.A. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Institutes of Health or the National Science Foundation.

Conflict of Interest Statement

The authors declare no conflict of interest.

CHAPTER 2 INTRODUCTION

While Chapter 1 focused on behavioral, morphological, and nutritional adaptations to plants, both Chapter 2 and Chapter 3 focus on the need for insect herbivores to overcome the toxic chemicals that plants create to defend themselves. Chapter 2 focuses on a way that herbivores may do this in the initial stages of evolving the ability to use plants–-through interactions with microbial symbionts in their gut. Bacteria that interact with plants also evolve ways to detoxify their chemicals, which in addition to being anti-herbivore are also anti-microbial. These adaptive traits of bacteria can be co-opted by their insect hosts. The intended outcome of this study was to characterize the gut community in *Scaptomyza flava,* especially focusing on bacteria that are able to detoxify plant defensive chemicals, and test whether these bacterial adaptations can also lead to trophic shifts in their hosts.

Characterizing the gut microbiome and its role in the evolution of herbivory in *Scaptomyza flava*

Jessica M. Aguilar*[1], Rebecca P. Duncan*[1], Diler Haji[1], Sneha Agarwal[1], Maria Mendoza[1], Kannagi Yashroy[1], Shivanni Sundram[1], Noah K. Whiteman[1,2]

*these authors contributed equally

1. Department of Integrative Biology, University of California, Berkeley

2. Department of Molecular and Cell Biology, University of California, Berkeley

Abstract

Plants have evolved to defend themselves against herbivory, and for herbivores, evolving the ability to attack plants is challenging. This push and pull between trophic levels has spurred the diversification of both plants and insect herbivores. There is also a third axis to these interactions–the herbivore gut microbiome. The gut microbiome has been shown to functionally benefit their host by providing metabolizing services of both nutrients and defensive chemicals. The gut microbial facilitation hypothesis posits that gut microbes may contribute to early adaptive leaps in the evolution of herbivory. Here, we developed a system to begin to test this hypothesis empirically. We first characterized the gut microbiome of a model herbivore, *Scaptomyza flava*, a drosophilid leaf-mining fly that eats mustard plants in the family Brassicaceae, including the genetic model plant, *Arabidopsis thaliana*. We found that the microbial community in the gut was most influenced by the life-stage of the fly, with host species (either *S. flava* or sister species *S. montana*) also playing a role. We then focused on identifying bacteria that are able to detoxify the common plant defensive chemical phenethyl isothiocynate (PEITC) using the SaxA enzyme. We isolated a suite of diverse Proteobacteria from the *S. flava* gut that are both resistant to and able to detoxify PEITCs using SaxA. Lastly, we used a genetic approach to test whether and to what extent bacteria with *saxA* conferred benefits to non-herbivores when challenged with PEITC compared to bacteria without *saxA*. We found that the functional contribution of detoxification by one bacterial species was not enough to confer an advantage. Future directions look toward testing the gut microbial facilitation hypothesis in this powerful tripartite system in other contexts and with a representative gut community.

Introduction

The evolution of herbivory in insects is hypothesized to be a particularly challenging trophic shift because of the constantly evolving obstacles presented by plants as a dietary resource and habitat (Southwood, 1972; Wiens et al., 2015). However, once herbivory evolves it tends to facilitate adaptive radiations, and has resulted in increased diversification rates of both the herbivores and the plants that they attack (Futuyma & Agrawal, 2009a; Wiens et al., 2015). Evolving the ability to use plants as both food and habitat requires novel adaptations to deal with the physical, nutritional, and chemically challenging sustenance that they provide. These hurdles can

be overcome through adaptation of genes already present in the genome, genes donated by other lineages through horizontal gene transfer, and through services provided by symbiotic microbiota (Engel & Moran, 2013; Gloss et al., 2014; Peláez et al., 2023, 2022; Whiteman et al., 2012). The advantages microbes can provide include supplementing nutrient metabolism (Baumann et al., 1995) and mitigating toxins (Ceja-Navarro et al., 2015; Mason et al., 2014; Shukla & Beran, 2020).

Plant toxins are a major challenge that may be overcome through symbiotic relationships with gut bacteria in the initial stages of the evolution of herbivory. The gut microbial facilitation hypothesis, first described in the context of the mammalian gut, proposes that microbes can provide herbivores an advantage by detoxifying plant defensive chemicals before endogenous detoxification methods are fixed in a population (Berenbaum, 1988; Douglas, 2013; Freeland & Janzen, 1974; Hammer & Bowers, 2015; Hansen & Moran, 2014; Wei et al., 2020). However, understanding the effects of the microbiome on the early stages of herbivory is challenging because interactions at early stages of this trophic shift are likely ephemeral. Bacteria are likely picked up from plants, and functional services that may be provided depend on the incidental community composition (Mason et al., 2014).

Here, we explored the interface between the herbivorous fly, *Scaptomyza flava*, the plants they use as hosts, and the microbes that live in their gut. *S. flava* is nested phylogenetically within the *Drosophila but* has evolved the ability to utilize plants as food and habitat. Ancestors of this species evolved an herbivorous lifestyle approximately 10-15 million years ago (Whiteman et al., 2012). *S. flava* has evolved various morphological, metabolic, and behavioral adaptations to utilize mustard plants in the family Brassicaceae as food, but they are still subject to the counter adaptations of their host plants (Whiteman et al., 2011). For example, mustard plants deploy isothiocyanates (ITCs) (among other metabolites and strategies) to defend against microbial pathogens and herbivores. ITCs are created in damaged leaf cells when their non-toxic precursors, glucosinolates, become hydrolyzed by myrosinases, which transform the molecule into a hydrophobic and electrophilic plant defensive compound (Figure 1a, Bones & Rossiter, 1996; Textor & Gershenzon, 2009). While there are several other insect systems with microbial symbionts that detoxify ITCs (Shukla & Beran, 2020; Welte, de Graaf, et al., 2016), the relatively recent transition to herbivory, well annotated genome, amenability to laboratory manipulations, and close phylogenetic relationship to the model fly *Drosophila melanogaster,* all make the *S. flava* system a well suited candidate to further explore the role the microbiome plays in insect interactions with toxic plants. *S. flava* have endogenous enzymes that protect them from ITCs, but their associated bacteria could also play a role in detoxification. Bacteria that utilize ITC-containing plants as hosts have also evolved ways to combat these antimicrobial plant defenses. One well described enzyme, SaxA, underlies detoxification of ITCs in many plant associated bacteria (Fan et al., 2011). Named by the conferral of the ability of bacteria to survive on *Arabidopsis* extract, *saxA* has been characterized as a plant virulence factor in bacteria and fungi that harbor this gene family (J. Chen et al., 2020; van den Bosch & Welte, 2017). Bacteria with genes in the *sax* family are able to metabolize ITCs into less toxic compounds by hydrolyzing the molecule and releasing a carbonyl sulfide that is further processed into hydrogen sulfide and carbon dioxide

(Figure 1b; Fan et al., 2011; van den Bosch et al., 2018; van den Bosch & Welte, 2017; Welte, de Graaf, et al., 2016; Welte, Rosengarten, et al., 2016).

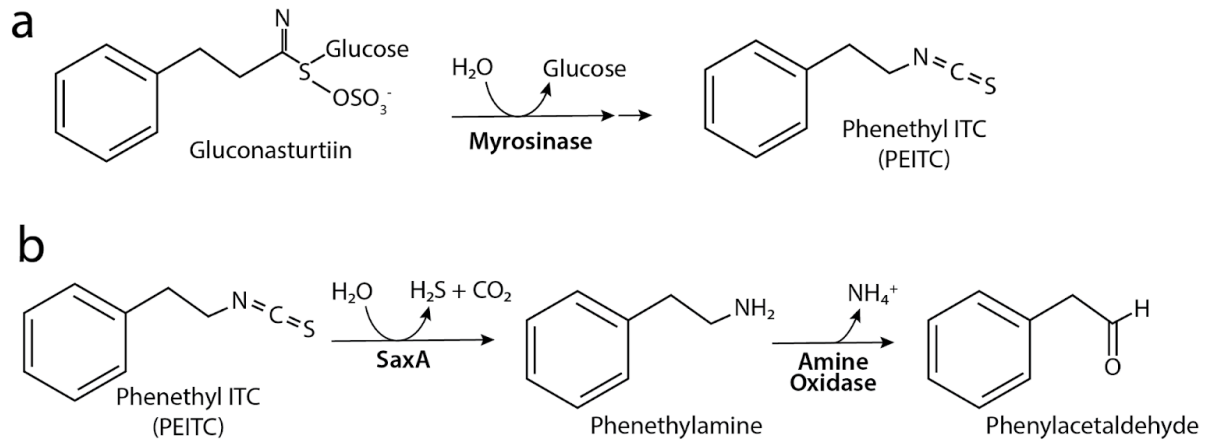


Figure 1a : Myrosinase activates the non-toxic glucosinolate. The glucosinolate precursor of PEITC, gluconasturtiin, is metabolized by the enzyme myrosinase and becomes a hydrophobic and electrophilic plant defensive compound. **Figure 1b: *saxA* enzymes hydrolyze PEITC by first removing carbonyl sulfide**. The remaining molecules are further processed into hydrogen sulfide and carbon dioxide and the less toxic phenethylamine and phenylacetaldehyde.

In this study, we first characterized the community of bacteria that live in the gut of the herbivorous drosphilid fly, *S. flava* in both larvae and adults. We then focused on identifying bacterial community members with the ability to resist and metabolize plant produced phenethylisothiocynate (PEITC), which is a highly abundant ITC produced during wounding in many different host plants of *S. flava*. We tested the gut microbial facilitation hypothesis–that the gut microbiota of nascent herbivores can facilitate the initial switch to herbivory by providing detoxification services of plant defensive chemicals in this system. We then used toxin feeding assays to understand aspects of the protective role of bacteria that are able to detoxify mustard oils in non-herbivorous flies.

Materials and Methods

Scaptomyza flava *collection, cultivation, and identification*

Laboratory colonies of *S. flava* used in this study originated from either a single female collected from watercress (*Nasturtium officinale*) growing in Strawberry Creek on the campus of the University of California, Berkeley, or from an outbred population collected in New Hampshire (Peláez et al., 2023). We reared lab colonies on *Arabidopsis thaliana* (Col-0) at approximately 22° C under a photoperiodicity of 16:8 (L:D). We sampled age structured, lab-reared larvae by extracting from leaves at 9-10 days after eggs were laid in plants. To collect eggs, we put lab-reared adult female flies (7-10 days after eclosion) in cages with *A. thaliana* (Col-0) and removed them after 48

hours. We sampled lab-reared adults and larvae from at least five independent rearing cages.

We collected the wild flies used in this study by visually searching leaves for mines (larval collection) or by sweep-netting patches of *N. officionale* (adult female collection) at UC-Berkeley. Wild-collected flies and larvae were either *S. flava* or the closely related, sympatric species *S. montana,* which cannot be distinguished morphologically in the larval phase; we further identified larvae and adults by a combination of morphological characteristics and the mitochondrial marker *cytochrome oxidase I* (*COI*) using the primers and thermal cycling conditions listed in Supplementary Table 1. We confirmed the phylogenetic position of sampled flies by inferring a phylogeny of *COI* using sequences of sampled flies and reference sequences from previously published reports (Peláez et al., 2023). We aligned sequences in MAFFT v.7.453 (Katoh & Standley, 2013) using default parameters and used for substitution model selection by the Akaike Information Criterion in jModeltest v.2.1.10 (Darriba et al., 2012). We inferred Maximum Likelihood phylogenies in RAxML v.8.2.12 (Stamatakis, 2014) on the CIPRES Science Gateway (Miller et al., 2011) and used the GTR+G (GTRGAMMA) model of nucleotide substitution. We preserved flies and larvae used for culture-independent analyses in 95% ethanol for DNA extraction at a later date.

Culture-independent community composition analysis

To assess the community composition of the gut-microbiome of lab-reared and wild-collected flies and larvae in a culture-independent manner, we homogenized individual preserved flies or larvae in sterile PBS and extracted DNA from the homogenates using Quick-DNA Fungal/Bacterial Miniprep kits (Zymo Research) following the manufacturer's protocol. There were too many samples to process at the same time, so to ensure consistency in how samples were treated, we included flies and/or larvae from different sampling locations/dates in each DNA extraction session. Finally, we performed at least one blank extraction (sterile PBS alone) for each extraction kit used to serve as contamination controls.

We sent DNA templates to the QB3 Sequencing Core at the University of California, Berkeley for 16S amplicon library preparation and sequencing on the Illumina MiSeq 300PE v3 platform. We amplified the 16S V3-V4 region with primers 515F and 806R, and mitochondrial and used plastid peptide nucleic acids (PNAs - see Appendix 2: Supplementary Table 1) to block background amplification of mitochondrial DNA and chloroplast DNA, which is present in consumed plant material present in fly guts. We also included PCR products from no template controls and a microbial community standard (Zymo, Catalog no. D6300) in amplicon sequencing.

16s rRNA Amplicon Data Processing

We sought to understand the factors influencing the S. flava gut microbiome community composition. To do this, we first denoised and merged reverse and forward

reads in QIIME 2 (Bolyen et al. 2018) using the DADA2 pipeline (Callahan et al. 2016) and then aligned the denoised sequences in MAFFT (Katoh and Standley 2013), filtered the alignments, and constructed a midpoint-rooted phylogeny using the "align-to-tree-mafft-fasttree" pipeline in QIIME 2. We classified ASVs with the QIIME 2 "classify-sklearn" plug-in after training a classifier with the "fit-classifier-naive-bayes" plug-in on sequences of the V4 region of 16s rRNA extracted from SILVA 99% OTUs (Release 132) database. The resulting feature tables, phylogenies, classifications, and sample metadata were primarily processed using the R package *phyloseq* (McMurdie and Holmes 2013). We removed ASVs classified to mitochondria, chloroplast, eukaryote, and archaea. Additionally, we removed ASVs that could not be classified at the Kingdom, Phylum, or Class taxonomic levels and ASVs with a total abundance across all samples within a dataset of less than three. We included only samples with a total abundance greater than 100. The remaining ASVs per dataset were further collapsed using the "tip_glom" function in *phyloseq* to cluster ASVs based on cophenetic distances with a tree height of 0.03. We then used the R package *decontam* (Davis et al. 2018) to identify putative contaminants using the prevalence-based filter with a threshold of 0.5 to remove ASVs that were enriched in the controls.

Culture-based bacterial isolation and community size measurements

To better understand the subset of the community of bacteria that can be cultured and used in downstream experiments, we isolated bacteria from lab-reared and wild-collected flies and larvae the same day as collection. We isolated bacterial colonies by homogenizing individual larvae or adult females in sterile PBS and spreading 10-fold serial dilutions (up to 1:1000 for larvae, up to 1:10,000 for adults) of fly homogenates on 10% Tryptic Soy Agar (TSA) medium. To prevent mold growth when isolating bacteria from adult flies, we supplemented 10% TSA medium with 2.88 mg/ml food grade natamycin. We incubated plates at 28°C for four days. For community size measurements, we extrapolated the total number of colony forming units (CFU) per fly from the number of visible colonies resulting from the dilution that yielded ~50-500 colonies. We randomly selected 50 colonies from each fly, stored them in glycerol stocks at -80°C and genotyped them for downstream assays. Some flies hosted very small communities with less than 50 colonies growing from the undiluted homogenate. In these cases, we selected all colonies for storage, genotyping, and downstream assays.

Assaying isolates for resistance to phenethyl isothiocyanate

We assayed up to 50 randomly selected isolates per fly for resistance to the commercially available aromatic isothiocyanate called 2-phenethyl isothiocyanate (PEITC) (Sigma), a common isothiocyanate and the most abundant isothiocyanate in the host plant of wild-collected flies. To determine resistance, we patched isolates onto freshly made TSA medium supplemented with 1% DMSO (solvent control) and TSA with 1% DMSO and 1500 μM PEITC. Plates were incubated overnight at 28°C. We scored isolates showing any growth on PEITC medium as resistant to PEITC.

Genotyping, collapsing, and classifying representative isolates

We genotyped the isolates assayed for PEITC resistance from haphazardly selected colonies from 4 lab-reared larvae, 4 wild-collected larvae, and 4 wild-collected adults (187 total isolates) by PCR-amplifying and sequencing full-length 16S using universal eubacterial primers 27F and 1492R (Appendix 2: Supplemetary Table 1). We trimmed low-quality ends of resulting sequences in Geneious (v.10.0.9) (https://www.geneious.com) using an error probability limit of 0.01. We clustered trimmed sequences into representative sequences in CD-HIT-EST (v.4.6) (Huang et al., 2010) using a sequence identity threshold of 1.0. We selected the longest sequence from each cluster as a representative sequence, and we classified bacterial isolates from which these sequences were obtained in the Ribosomal Database Project Classifier (https://rdp.cme.msu.edu/classifier/classifier.jsp).

Assaying resistant isolates for PEITC metabolism

We assayed select representative isolates showing resistance to PEITC for the ability to metabolize the compound, focusing on isolates from the bacterial family Enterobacteriaceae and the genus *Pseudomonas* in the Pseudomonaceae. Enterobacteriaceae and *Pseudomonas* are common environmental bacteria that were also abundant in our culture-independent analysis of community composition in both lab-reared and wild-collected flies (Figure 2a). In total, we assayed 61 isolates representing taxa that were found in both lab-reared and wild-collected flies. To determine if isolates could metabolize PEITC, we used a colorimetric assay (Cline assay, (Cline, 1969)) that detects bacterial production of hydrogen sulfide (H_2S), a product of bacterial isothiocyanate breakdown (Figure 1) (Tang et al., 1972; Welte, de Graaf, et al., 2016), through methylene blue formation (Cline, 1969; Engelbrektson et al., 2014).

We first used a rough screen to detect H_2S production in 10 focal isolates. We grew isolates overnight at 28°C in tryptic soy broth (TSB), they were then pelleted, washed three times in PBS, and resuspended in PBS. We incubated bacterial suspensions in a 96-well plate in a shaking incubator for 22 hours with or without 100 mM PEITC. Following incubation, we pelleted bacteria and added equal volumes of Cline reagent A (10 mL Sulfuric acid 0.5 g Zinc acetate dihydrate, 0.5 g N,N dimethyl-p-phenylene-diamine-sulfate, filled with ddH_2O to a final volume of 500 mL), sample supernatant, and Cline reagent B (2.5 g $FeCl_3$ * $6H_2O$, 2.5 g $FeCl_3 \cdot 6H_2O$, filled with ddH_2O to a final volume of 500 mL). We mixed and incubated the samples at room temperature for at least 10 minutes, then measured absorbance at 660 nm.

For each isolate, we subtracted the background absorbance of a sample without PEITC from absorbance of a sample with PEITC. We selected isolates for a more systematic assay with technical replication if they were visibly blue after adding Cline reagents A and B, or if the absorbance at 660 nm was at least twice the background. For the follow-up assay, we incubated concentrated suspensions of bacterial isolates in PBS (OD_{600} = 20) in triplicate with or without 100 mM PEITC in 96-well plates that we

sealed with aluminum sealing film. We tested *S. flava* isolates alongside several controls. Positive controls included a strain of *Pseudomonas syringae* (DC3000) known to metabolize PEITC (Fan et al., 2011; Welte, de Graaf, et al., 2016), and a strain of *Enterobacter cloacae* isolated from *C. elegans* (Berg et al., 2016) that encodes an isothiocyanate hydrolase (*saxA*). Negative controls included a standard laboratory *E. coli* strain (NEB DH-alpha) and a *saxAB* deletion mutant of DC3000 (Fan et al. 2011). We used PBS alone and PBS with PEITC (no bacteria) to determine background absorbance. We incubated plates as described for the rough screen, and following incubation we pelleted bacteria and the Cline assay was run as described. For each isolate and the PEITC without bacteria samples, absorbances (A_{660}) of supernatants mixed with Cline reagents (with background subtracted) were averaged across technical replicates and plotted as a proportion of the average A_{660} of DC3000, which was set to 1. Since these data were not normally distributed, we performed a two-way ANOVA on rank-transformed data, followed by a Tukey's Honestly Significant Difference (HSD) test in R to determine differences between treatments (R Core Team, 2021). We then used the cldList function in rcompanion (Mangiafico & Mangiafico, 2017) to identify groups of samples that did not significantly differ.

Bacterial genome sequencing, and assembly

To identify whether SaxA was involved in metabolizing PEITC, we purified genomic DNA from 20 pelleted isolates showing H_2S production in the Cline assay using a ZR Fungal/Bacterial DNA Miniprep kit (Zymo) following the manufacturer's protocol. Bacterial genomic DNA was submitted to the QB3 Genomics Core Facility at UC Berkeley for library preparation and paired-end DNAseq on the Illumina HiSeq 4000 platform. We checked read quality with FastQC (Brown et al., 2017). Draft genomes for isolates were assembled in SPAdes (v.3.13.0) (Bankevich et al., 2012) using default parameters.

saxA annotation and unique isolate validation

Although there may be other novel enzymes that confer resistance to PEITC, the enzyme SaxA can hydrolyze isothiocyanates and has been described in multiple PEITC metabolizing microbial strains (Fig. 1B) (Fan et al., 2011; Welte, de Graaf, et al., 2016). To determine if saxA might mediate PEITC metabolism in *S. flava* isolates, we used the seven functionally characterized saxA protein sequences (GenBank IDs: WP_012863804.1, WP_039508597.1, NP_791682.1, WP_005967032.1, KFF62654.1, ANJ99341.1, WP_094484734.1) (van den Bosch et al., 2018) as queries for TBLASTN searches against isolate genome assemblies. We confirmed scaffold regions hit by all seven queries with an e-value less than 0.001 to be *saxA* orthologs using a phylogenetic approach. The seven functionally characterized saxA proteins fall across four distinct monophyletic groups within the part of the *saxA* gene family that includes many sequences from Enterobacteriaceae and *Pseudomonas* (van den Bosch et al., 2018). We amassed a group of sequences representing this part of the *saxA* gene family by using the functionally characterized protein sequences as queries in BLASTP searches against NCBI reference proteins and selecting the top 10 hits of each query

for our phylogenetic analyses. We aligned protein sequences from BLASTP searches against the NCBI refseq database and TBLASTN searches against our isolate genomes in MAFFT v.7.453 (Katoh & Standley, 2013)) using default parameters and the alignment was run through ProtTest 3 (Darriba et al., 2011) to determine the best-fit model of amino acid substitution by the Akaike Information Criterion. We then inferred phylogenetic trees by Maximum Likelihood in RAxML (v.8.2.12) (Stamatakis, 2014).

Some isolates encoded identical *saxA* sequences, so we used a reciprocal best BLAST approach to annotate four molecular markers (*gyrB*, *rpoD*, *gapA*, and *gltA*) and determine whether isolates were unique. We subjected scaffolds from isolate genomes to TBLASTN searches using as a query for each gene the protein sequence from *E. coli* K-12 substr. MG1655 (taxid: 511145) (*gyrB*: YP_026241.1; *rpoD*: NP_417539.1; *gapA*: NP_416293.1; *gltA*: NP_415248.1). We extracted aligned sequences from scaffolds and collapsed isolates with identical nucleotide sequences into one representative, unique isolate. We then translated aligned sequences and subjected them to BLASTP searches against the protein refseq database for *E. coli* subst. MG1655. We used sequences that hit the original query sequence in BLASTP searches against the NCBI protein refseq database to confirm the taxonomic identity of each unique isolate.

Creation of CeEnt1 ΔsaxA *mutant*

To test whether the *saxA* gene was sufficient to confer a benefit to non-herbivorous flies when challenged with PEITCs, we created a *saxA* deletion mutant of the laboratory strain CeEnt1 using the λ recombinase system (Murphy, 1998) to use as a control. CeEnt1 is a lab *E. cloacae* strain isolated from *C. elegans* (Berg et al., 2016) that contains one copy of *saxA.* This strain metabolized PEITC at a rate comparable to the most highly efficient wild bacteria tested. To create the mutant, we transformed electrocompetent CeEnt1 with a pKD46 plasmid and a PCR product that had a kanamyacin resistance gene flanked with FRT sites and approximately 50 bp of sequence homologous to the region just before and after the *saxA* gene. The λ Red machinery was then used to perform homologous recombination between the flanking regions, and the kanamycin resistance gene replaced *saxA*. We then selected for cells containing kanamycin resistance and confirmed the deletion by sequencing. We then validated that the Δ*saxA* mutant lost function by comparing resistance between wild-type CeEnt1 and Δ*saxA* by plating 7-fold serial dilutions of each type of bacteria with starting, undiluted concentrations of OD600 = 1 on plates with various concentrations of PETIC (0, 6.25, 12.5, 25, 50, 100, and 200). We then extrapolated the total number of CFUs that were able to grow on the various concentrations from the number of visible colonies resulting from the dilution that yielded easily discernible colonies. We also tested the ability to metabolize PETIC in the cline assay described above.

In vivo *PEITC metabolism assay*

Next, we tested whether having bacteria with the ability to detoxify PEITCs in the gut extends the fitness of flies that are naive to these toxins using Capillary Feeding

(CaFe) assays. The CaFe assay vial consists of a small, clear plastic container (approximately 1 in diameter by 2 inches in height) with air holes drilled into the bottom. This container is placed into a larger container with water in the bottom to increase the humidity in the vial for the fly. The flies are offered food through 5 µl glass capillary tubes that are fed through well fitting holes in the lid. Because of their amenability to grow in lab culture and their relatedness to *S. flava*, we used the yeast feeding *D. melanogaster* in this assay. We first collected flies 0-3 days post emergence and placed 1 female into each CaFe assay vial (Ja et al., 2007). First, we presented flies a sterile antibiotic cocktail consisting of 50 µg/ml tetracycline, 100 µg/ml rifampicin and 100 µg/ml streptomycin in a 5% sucrose solution for 24 hours. Then, we allowed flies to feed on sterile 5% sucrose for two days before introducing bacteria for two days. The bacteria treatments we used were wild type CeEnt1, Δ*saxA*, and heat killed CeEnt1 that was boiled in a water bath for 5 minutes and cooled. Bacteria was first grown on LB plates, then cultured overnight in LB broth. We pelleted, then diluted bacterial suspensions to OD600 = 0.6 in 5% sucrose. We validated that the living bacterial treatments were able to persist in 5% sucrose and quantified the ability of these bacterial treatments to colonize the gut before beginning (Appendix 2: Supplementary Figure 3 and 4). Then, after allowing the flies to feed on bacterial treatments for two days, we began feeding PEITCs across a range of concentrations (0, 1.25, 2.5, 3.75, 5, 10) mixed in sterile 5% sucrose solution, vortexing between every fill. We replaced the capillary tubes daily and measured the amount removed from each tube. Flies were kept randomized in an incubator at 24 degrees Celsius and 60% humidity with a 12:12 light:dark cycle. We also included vials without flies in our randomization to estimate the rate of evaporation in each PEITC concentration. Replicates where evaporation controls exceeded or equaled the amount eaten were excluded. Since these data were not normally distributed, we performed a two-way ANOVA on rank-transformed data, followed by a Tukey's Honestly Significant Difference (HSD) test in R to determine differences between treatments (R Core Team, 2021).

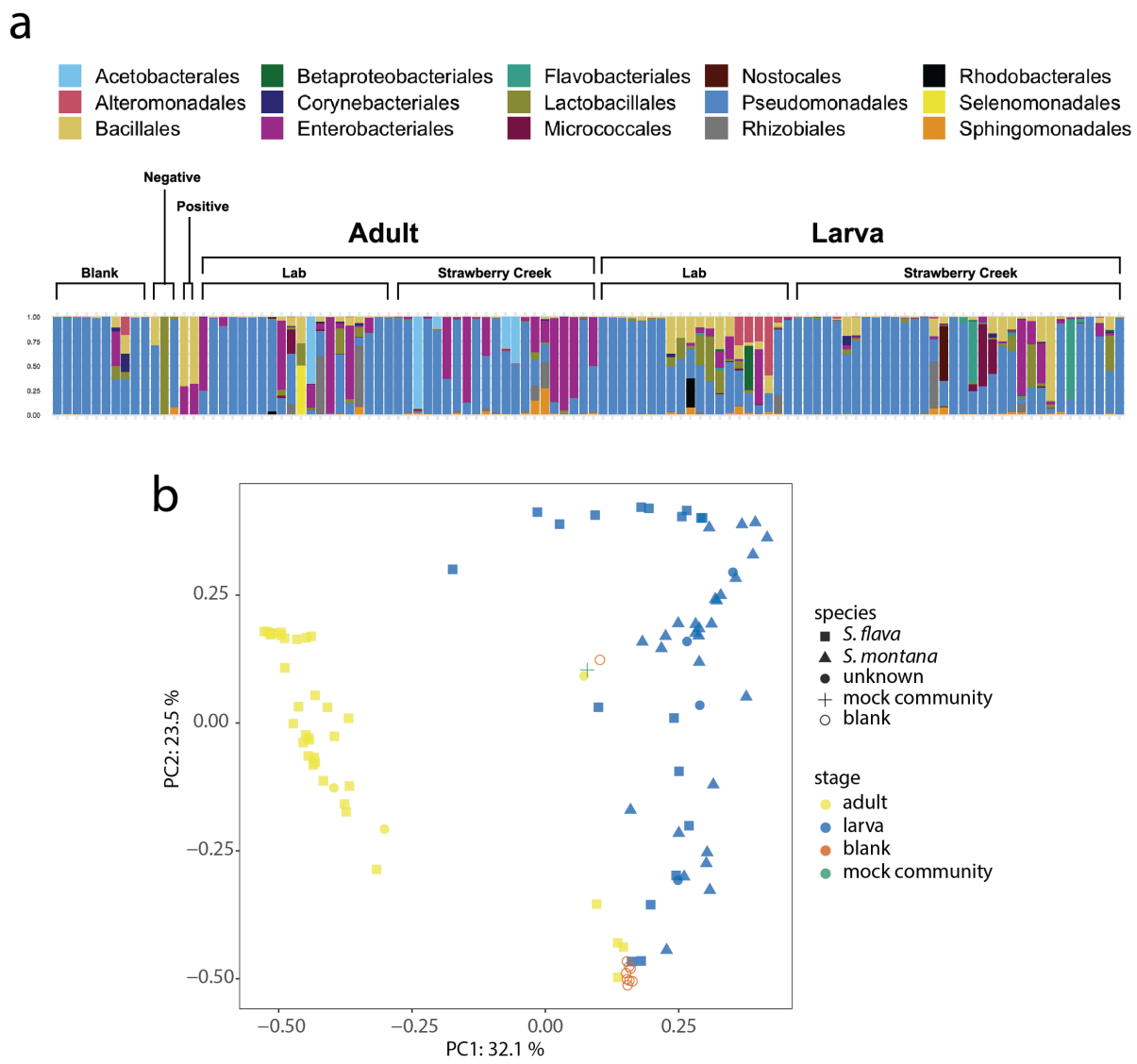


Figure 2a: *Scaptomyza* bacterial community composition. Relative abundance of top 12 most abundant bacterial orders across samples. Samples are labeled to indicate blank extractions of sterile PBS, negative controls (no PCR template), positive controls (microbial community standard), adult flies, or larval flies. Adult and larval fly samples are further labeled to indicate whether they were collected in Strawberry Creek in Berkeley, CA (21 adults, 34 larvae) or sampled from lab-reared populations (20 adults and 20 larvae). Wild larval populations are a mixture of *S. flava* and *S. montana*, and were identified using COI sequencing. **Figure 2b:** PCoA of Bray-Curtis distances using DADA2-denoised read counts across all samples. No additional filtering was performed.

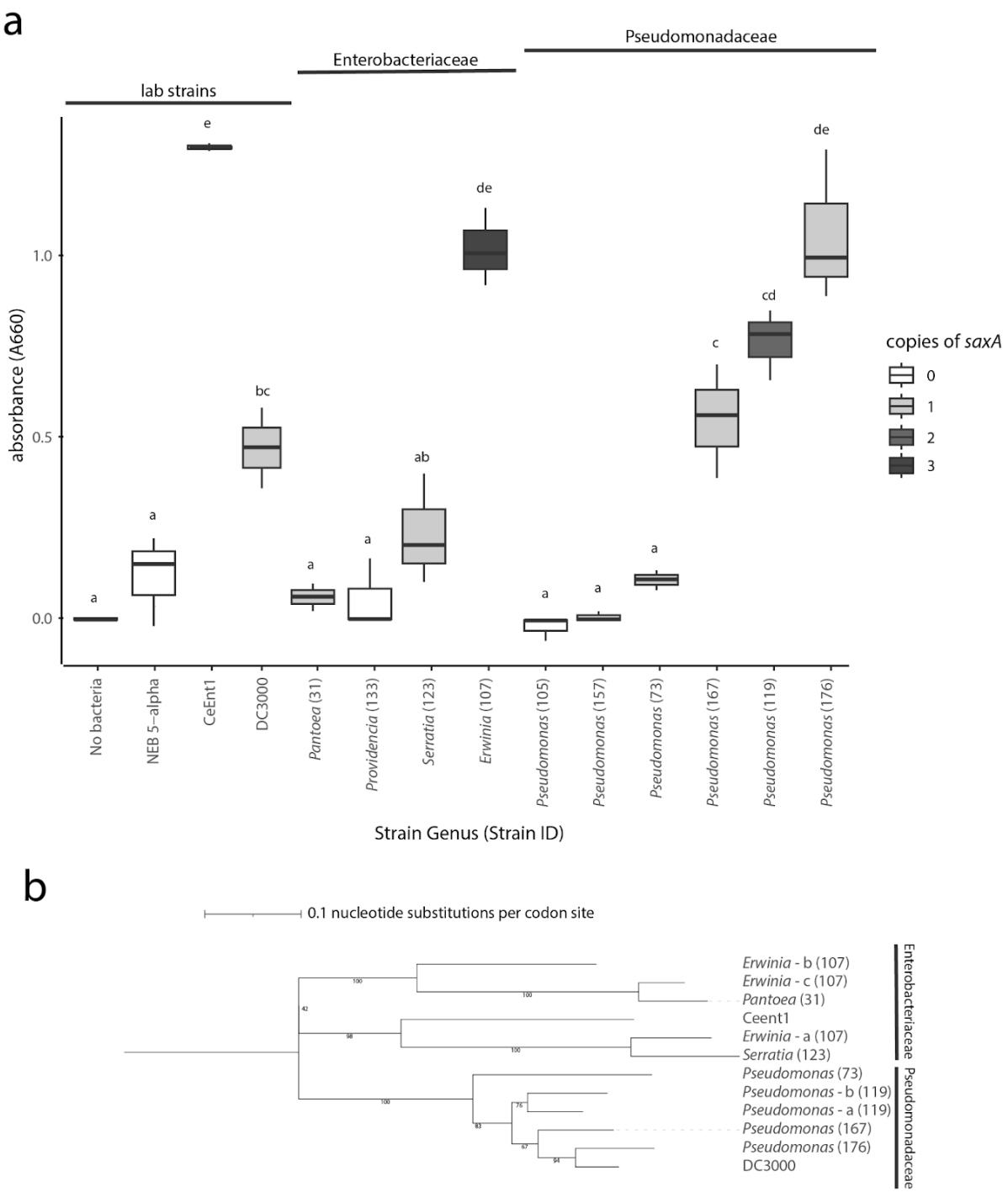


Figure 3a : **Absorbance measurements at 660nm, which indicates the amount of PEITC metabolized in the Cline assay.** Bar color indicates the number of copies of saxA present in each strain. Wild strains are characterized by Order and lab strains are indicated. CeEnt1 is in the order Enterobacteraceae and DC3000 is in the order Pseudomonadaceae. One copy of saxA is required to metabolize PEITC, but does not always confer the ability to metabolize. **Figure 3b** : **Phylogenetic relationships of *saxA* copies found in wild flies.** Nucleotide aware maximum-likelihood tree was created from *saxA* sequences found in bacteria in the gut of flies found in Strawberry Creek in Berkeley, CA. Tree is rooted using the rootstrap method in IQ-TREE followed by tree topology tests.

Results

Scaptomyza community composition is driven by lifestage

The goal of this work was to test the hypothesis that gut bacteria facilitate a transition to herbivory. To test this hypothesis, we first characterized the bacterial communities of the lab-reared and wild-collected herbivorous flies *S. flava* and *S. montana* to identify candidate taxa that may play a role in detoxification of their host plants. The most common bacteria in larvae and adults reared on *Arabidopsis* and collected from watercress were in Pseudomonadales and Enterobacteriales (Figure 2a). After plotting the first two principal components in a principal component analysis of Bray-Curtis dissimilarity, we found that life stage seemed to drive most differences (Figure 2b). Species also contributed to differences, but there was overlap in bacterial composition in some individuals of different species. Collection date did not drive differences in community composition (Appendix 2: Supplementary Figure 1).

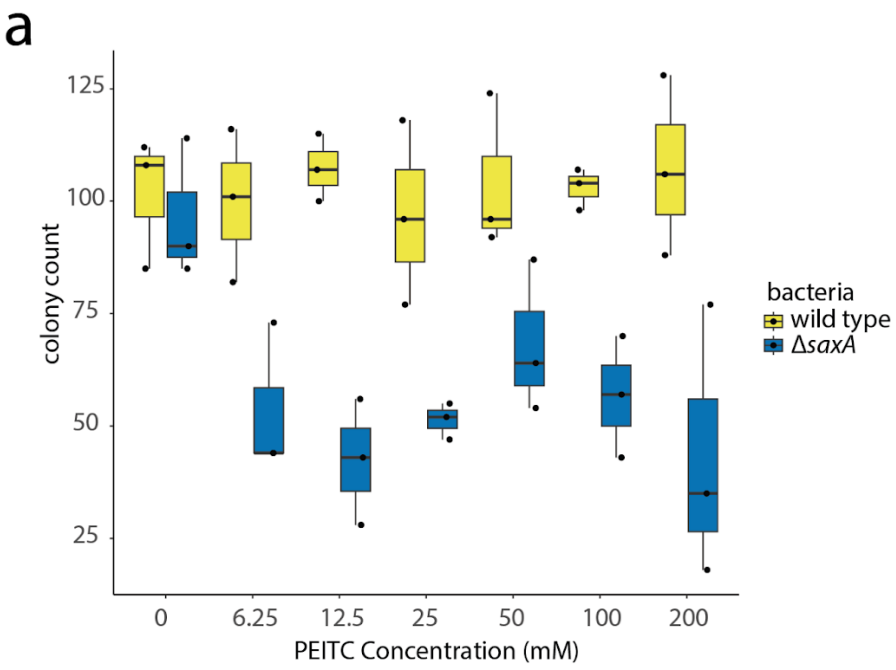
Pseudomonas and Enterobacteriacae isolates from lab and wild metabolize PEITC

To determine whether bacterial strains isolated from the gut of *S. flava* could metabolize PEITC, we first assayed wild collected strains for resistance on PEITC through their ability to grow in PEITC-containing plates. We then further tested 61 isolates representing taxa that were found in both lab-reared and wild-collected flies for PEITC metabolization. In a Cline assay, we measured the absorbance at 600 nm of a sample that has undergone a chemical reaction that creates various shades of blue, depending on the amount of hydrogen sulfide (H_2S) present (Cline, 1969; Engelbrektson et al., 2014). The amount of this compound is proportional to the ability of a particular bacteria to metabolize PEITC into less toxic forms, phenethylamine and phenylacetaldehyde (Figure 1b). We measured the absorbance of strains that we found to be resistant to PEITC and NEB 5-alpha, which was used as a negative control (Figure 3a). We found differences in PEITC metabolism explained by strain ($f = 4.7824$, $p = 0.002$). Although at least one copy of *saxA* was necessary to surpass the level of absorbance in the negative control (NEB 5-alpha), having a copy of *saxA* was not sufficient in conveying the ability to metabolize PEITC (i.e. *Pantoea* strain 31, *Pseudomonas* strain 157, and *Pseudomonas* strain 73). Phylogenetically, we found that *saxA* copies are clustered among taxonomic groups, with *Enterobacter* and *Pseudomonas* clustering together, respectively (Figure 3b).

saxA *deletion mutant does not metabolize PEITC*

After creating a *saxA* deletion mutant in *Enterobacter cloacae* Ceent1 strain (Δ*saxA*), we validated its ability to resist and metabolize PEITC. We first performed a resistance assay that compared the ability of wild type CeEnt1 and Δ*saxA* to grow on PEITC-containing media (Figure 4a). We found that Δ*saxA* mutants were less resistant to PEITC ($f = 61.229$, $p = 4.15e^{-9}$) regardless of concentration ($f = 1.468$, $p = 0.218$), but were still able to grow. We then validated that Δ*saxA* mutants were unable to metabolize PEITC in a Cline assay (Figure 4b). We found that the ability of our Δ*saxA*

strain to metabolize PEITC was completely removed (p = 0.007). We tested whether our experimental bacteria strains colonized the gut with a concentration of OD600 = 0.6, and found bacteria present in most flies tested after one day, with variable amounts of persistence (Appendix 2: Supplementary Figures 3 and 4). We proceeded with this concentration, as we found that higher concentrations of initial inocula were harmful to fly health. Additionally, flies in the wild will be consistently inoculating their gut bacteria, and a boost in fitness that may be present for one day might still be relevant over the course of a four-day experiment.



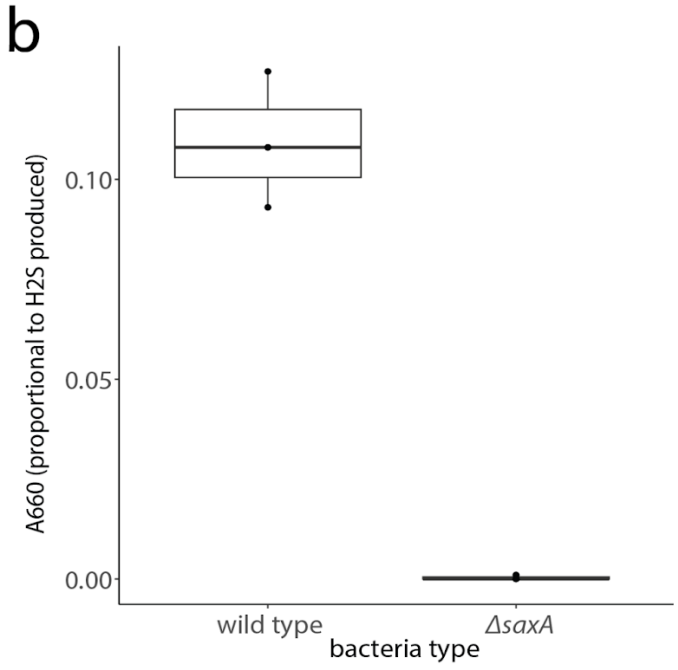


Figure 4a : Experimental bacteria resistance assay to ITCs. We found the growth of Δ*saxA* was reduced compared to wild type CeEnt1 on plates with increasing amounts of PEITC. **Figure 4b**: **Experimental bacteria Cline assay.** Metabolism of PEITC was eliminated in Δ*saxA* bacteria. Welch Two Sample test, p-value = 0.007992.

saxA *alone does not confer an advantage on PEITC*

To determine whether having PEITC metabolizing bacteria in the gut microbiome was sufficient to convey a fitness benefit to flies eating a PEITC rich diet, we performed a CaFe assay where the non-herbivorous *D. melanogaster* were first inoculated with wild type CeEnt1, Δ*saxA,* and heat killed CeEnt1 then fed toxins at various concentrations. We found that the amount of food flies ate depended on whether there was PEITC in their food, with flies eating less with PEITC present (Figure 5b; $f = 8.27$ $p = 2.28e^{-7}$). There was no difference in the amount of PEITC solution eaten between PEITC concentrations or bacteria type in the gut (Appendix 2: Supplementary Table 1). Additionally, there was a difference in the dose (µg/fly) of toxin eaten (Figure 5c; $f = 107.0$ $p < 2e^{-16}$); flies that ate the highest concentrations of toxin were exposed to a significantly higher dose of PEITC (Appendix 2: Supplementary Table 2). Finally, we compared the response curves between bacteria types. We found that bacteria genotype did not have an effect on the survival of flies on various concentrations of PEITC during the duration of the experiment (Figure 5d, $\chi^2 = 0$, $df = 2$, $p = 1$, Appendix 2: Supplementary Figures 5-7).

Discussion

The evolution of herbivory in insects has spurred diversification in both the insects and the plants that they use as hosts (Ehrlich & Raven, 1964; Futuyma & Agrawal, 2009a). There are hurdles presented by plants that herbivores must overcome, especially when plants are also evolving counter defenses (Mitter et al., 1988; Schoonhoven et al., 2005; Southwood, 1972). Herbivores have evolved counter-strategies to meet the challenges posed by plants, either through endogenous evolution or sometimes through symbiotic relationships with microbes in their gut that provide functional services for their insect hosts.

Here, we first characterized the community of microbes in the gut of lab and field-collected individuals of the herbivorous fly, *S. flava*. We found that gut bacterial community composition was variable between individual flies (Figure 2a), with developmental stage driving the most consistent differences in bacterial composition and species driving differences in some cases (Figure 2b, Appendix 2: Supplementary figure 1). Plant and larval microbial communities overlap significantly (O'Connor et al., 2014), so differences in community composition between adults and larvae are likely due to the differences in the interactions with plants between these two life stages. Although larvae can move from leaf to leaf, they are more restricted in their interactions with plants when compared to adult flies which interact much more with the exterior structures of the leaf, especially through holes made in the leaf surface by females with their serrated ovipositor (Peláez et al., 2022). Second and third instar larvae can move between leaves if their food source is depleted, but they generally stay within the mesophyll, scrape the tissue with their enlarged mouth hooks and consume the plant fluids that are released (Aguilar et al., 2024). Adults and larvae are both exposed to the toxic chemicals produced by their damaged plant hosts.

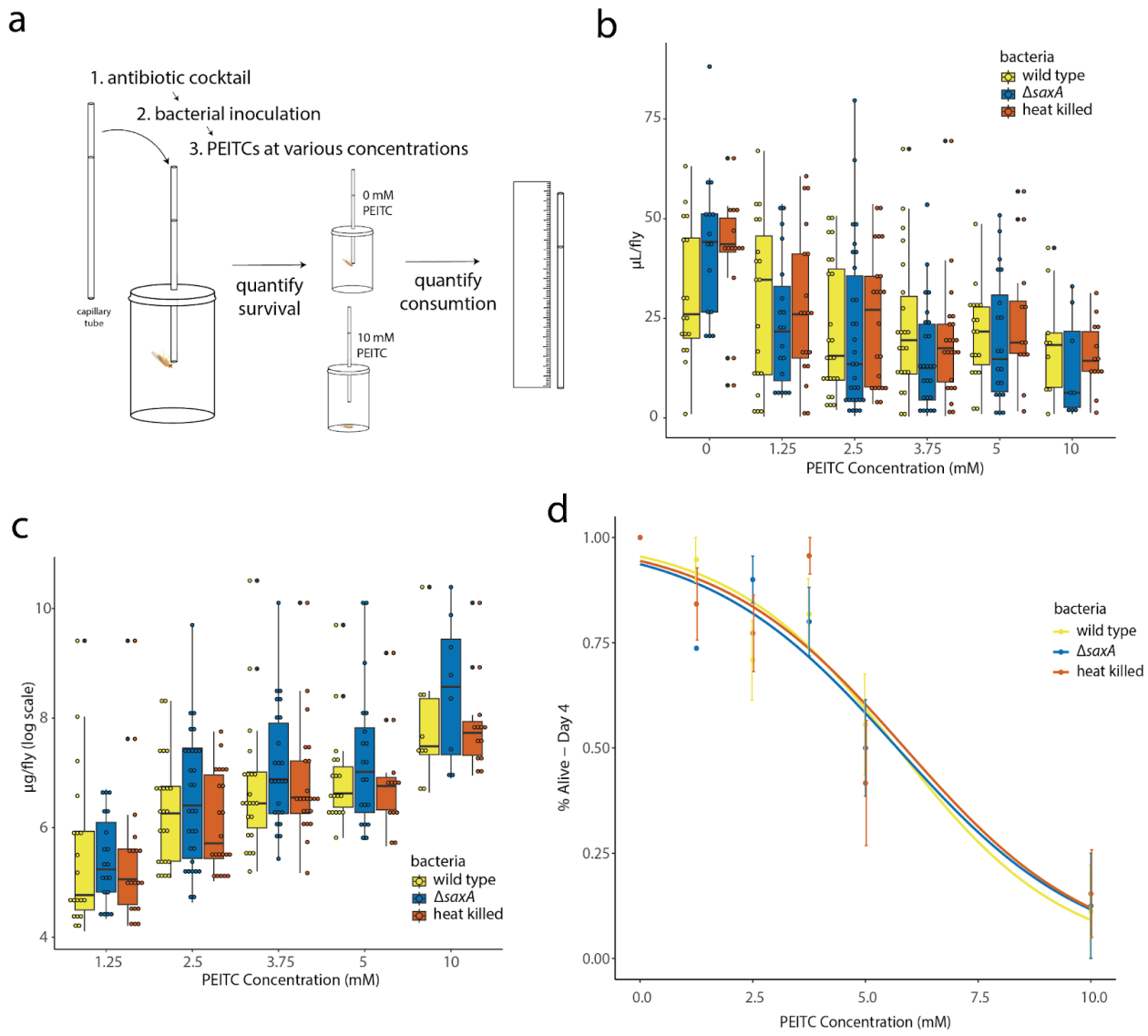


Figure 5a : **Experimental set up for the CaFe assay.** We used capillary tubes to feed D. melanogaster sucrose mixed with either antibiotics, bacteria (to re-inoculate their gut), or PEITC. We then quantified survival and consumption. **Figure 5b : Amount of PEITC in 5% sucrose eaten.** We found the PEITC solution to be generally aversive to feeding for *D. melanogaster* across bacteria types. **Figure 5c** : **Amount of PEITC consumed by *D. melanogaster*, in µg.** Across bacteria types, flies in the highest concentration consumed the largest amount of toxin. **Figure 5d : CaFe Assay.** We cleared the gut of *D. melanogaster* of bacteria, then re-inoculated with wild type CeEnt1 (that is able to metabolize PEITC) and Δ*saxA* (that can not metabolize PEITC). We then fed flies PEITC at various concentrations and tracked survival.

Generally, a main challenge for species that interact with mustard plants is overcoming their characteristic defensive chemicals, especially ITCs. Glucosinolates can be divided into three categories, aliphatic, indolic and aromatic, depending on the amino acid precursor of the compound. Although nontoxic themselves, when plant tissues are damaged glucosinolates are activated by the enzyme myrosinase and become various toxic, plant-protecting compounds, like ITCs in the case of aliphatic glucosinolates. Bacteria that live on or in mustard plants detoxify ITCs with an enzyme that hydrolyzes them called SaxA (Fan et al., 2011; van den Bosch et al., 2018). We characterized the ability of *S. flava* gut bacteria to resist ITCs by their ability to grow on media with PEITC (a common ITC present in the host plants of *S. flava*) and metabolize this defensive chemical by measuring the amount of hydrogen sulfide that is released when bacterial strains and PEITC are incubated together in a Cline assay (Figure 3a (Cline, 1969). We also searched the genome of resistant strains to see if *saxA* was present. In the strains we tested, having at least one copy of *saxA* was required for metabolizing PEITC, but detoxification did not scale with the number of copies of *saxA* present in the bacterial genome. The *E. cloacae* (CeEnt1) that we used in subsequent experiments has one copy of *saxA* and was among the most efficient at detoxifying PEITC. A different *E. cloacea* strain isolated from *Carica papaya*, in the Brassicales, was the first bacterium shown to be able to detoxify an ITC (benzyl ITC) using a hydrolase to produce hydrogen sulfide gas and an amine (Tang et al., 1972). Additionally, some resistant strains with *saxA* were unable to detoxify PEITCs, indicating that there are likely other, non-*sax* mechanisms that don't rely on removing hydrogen sulfide to disarm PEITC that allow bacteria to live in the presence of PEITC. Detoxification of ITCs by *saxA* has been implicated in several other systems. Bacteria and other microbes evolve the ability to detoxify ITCs through their own interactions with plants, and *saxA* is implicated as a virulence factor for these organisms (van den Bosch et al., 2020). In addition to the bacteria that we focused on, there is also a documented case of a plant pathogenic fungus, *Sclerotinia sclerotiorum*, that utilizes *saxA* to detoxify ITCs and promote virulence (J. Chen et al., 2020).

S. flava has the potential to be an informative model system for herbivore-microbe interactions, especially in light of the knowledge and tools that come with attacking the model plant, *A. thaliana* and being closely related to the microbe feeding, non-herbivorous model organism, *D. melanogaster*. In addition to its status as a premier genetic model animal, *D. melanogaster* is also considered a model for understanding the gut microbiome (Ludington & Ja, 2020). For both *S. flava* and *D. melanogaster*, differences in fly habitat are implicated in shaping the structure and function of the bacterial community (Bost et al., 2018; O'Connor et al., 2014). There are clear nutritional and metabolic benefits of the microbiome for *D. melanogaster* (Grenier & Leulier, 2020). Related to our goals here, experimental evolution studies have shown the potential for using *D. melanogaster* as an experimental host species to further our understanding of the role the microbiome plays in adaptations to dietary toxins, like the adaptations to toxic noni fruit utilized by *D. sechellia* (Heys et al., 2021). For *S. flava*, the presence of bacterial strains in the gut that can metabolize diet-derived ITC may also play a role in host fitness and the results of our community analysis indicate that there are several strains that metabolize these plant compounds (Figure 3a; (Shukla & Beran, 2020)).

In the earliest stages of the evolution of herbivory, relationships with gut bacteria may provide assistance in dealing with plant toxins before endogenous adaptations evolve–this is called the gut microbial facilitation hypothesis (Freeland & Janzen, 1974; Hammer & Bowers, 2015). We tested this idea in *D. melanogaster*. We replaced the native microbiome of these flies with wild-type bacteria that have *saxA* or have that gene deleted and so are unable to detoxify PEITC (Figure 4a and 4b, Appendix 2: Supplementary Figures 3 and 4). We then fed these flies a range of PEITC concentrations, quantified the amount they ate, and quantified their lifespan. We found that PEITC was non-appetitive when present, regardless of concentration or of bacterial treatment, though dose did increase with increasing concentrations. This aversion is likely due to *D. melanogaster*'s well characterized gustatory aversion to bitter compounds (like ITCs) (Kang et al., 2011; Lichtenstein et al., 1962; K. Scott, 2018). Generally, interactions between insects and their microbiomes are highly variable and context dependent, spanning ephemeral inoculation from feeding on plants to highly specialized co-dependencies (Baumann et al., 1995; Mason, 2020; Mason et al., 2014; O'Connor et al., 2014). In the context of single species inoculations with the ability or inability of detoxification through metabolism by *saxA,* survival outcomes were nearly identical, but there still may be some benefit to having PEITC degrading bacteria in a gut community context or when insects are interacting with more specialized gut commensals. Bacteria that are unculturable or that utilize detoxification methods outside of *saxA*-mediated hydrolysis may also play a role. More study is needed to determine whether the gut microbiome may play a role in the evolution of herbivory.

In conclusion, we have laid the groundwork for utilizing *S. flava* and its associated gut microbial community as a model for understanding tripartite plant-microbe-herbivore interactions. We have characterized the community structure of the gut microbiome of the herbivore *S. flava* in the lab and in a wild population. We adapted powerful tools, including the CaFe and Cline assays and leveraged the close relationship of *S. flava* with model species, *D. melanogaster.* In future studies, we are primed to dig deeper into the nuances of the potential benefits and costs of bacterial community interactions with their insect hosts in laboratory settings varying from CaFe vials to living plants and in the various natural habitats of *S. flava* in the northern hemisphere. Understanding these tripartite interactions will allow us to better understand the role the microbiome might play in the evolution of herbivory.

CHAPTER 3 INTRODUCTION

Like Chapter 2, Chapter 3 focuses on the need for herbivores to detoxify plant defensive chemicals. Since the mid 20th century, herbivores and other insects have also needed to withstand synthetic pesticides–the increase of their use in modern agricultural practices has placed a huge selective pressure on insects. However, insect herbivores, with their pre-adaptations to toxic plant hosts, are primed to overcome these novel xenobiotics. We focused on a gene well characterized in its ability to detoxify insecticides, *cyp6g1*. The herbivorous *Scaptomyza flava* has an impressive seven functional copies of this gene; these duplication events happened long before the advent of pesticide use. Here we studied the evolution of *cyp6g1* throughout herbivorous and non-herbivorous *Scaptomyza,* and sought to understand the roles that adaptive evolution, gene duplication, and transposable elements play in shaping *cyp6g1* copies.

The evolution of the xenobiotic detoxification gene *cyp6g1* in the herbivorous *Scaptomyza*

Jessica M. Aguilar[1], Josephine Tamsil[3], Rebecca L. Tarnopol[2], Susan L. Bernstein[1,3], and Noah K. Whiteman[1, 3]

1. Department of Integrative Biology, University of California, Berkeley

2. Department of Plant and Microbial Biology, University of California, Berkeley

3. Department of Molecular and Cell Biology, University of California, Berkeley

Abstract

A major goal of evolutionary biology is to understand the factors that underlie adaptation to the biotic and abiotic environment. Herbivorous insects have provided important evidence for the mechanisms driving adaptation, particularly through biotic interactions with their host plants and human-mediated interactions with insecticides. The adaptations underlying detoxification of phytotoxins and synthetic pesticides overlap in herbivorous insects. The pre-adaptation hypothesis posits that the ability to detoxify some plant chemicals may lead to higher likelihood of successful metabolizing synthetic pesticides, but this hypothesis is still relatively untested. *Cyp6g1* encodes a Cytochrome P450 monooxygenase that has been implicated in both insecticide resistance and phytochemical detoxification in several insect species. Here, we compared *cyp6g1* in related herbivorous and non-herbivorous drosophild species, and using a branch-site test, found signatures of positive selection along the branch leading to the bulk of the *cyp6g1* copies in herbivores. We then focused on *S. flava*, which has seven functional copies of *cyp6g1* and a P-transposon inserted between two of those copies. Although the impact of their presence is not always clear, transposable element insertions *cis* of detoxification genes tend to be related to an increase in the ability to detoxify. We studied the population genetics of this insertion in several pooled datasets and found that it was present in all three of the focal populations which spanned two sites across two years. To further understand the possible impact on expression of the P-transposon in relation to xenobiotics, we measured differential expression of *cyp6g1* copies in response to plant defensive chemicals. We found an increase in expression of one copy (*cyp6g1h*) in response to plant defensive chemicals. Though, the P-transposon is not directly *cis*, indicating it is unlikely to have influenced this increase. Finally, we developed a homologous expression assay to measure the affinity of each Cyp6g1 copy across the clade to detoxify either indole glucosinolate derived plant defensive compounds or insecticides. Further results of this assay are forthcoming.

Introduction

Herbivorous insects and the plants that host them are the two of the most diverse macroscopic groups of organisms on the planet. It is widely accepted that they have

become diverse because of their ecological interactions (Ehrlich & Raven, 1964; Labandeira & Currano, 2013). The ongoing push and pull between plants and their defensive chemicals and subsequent counter adaptations in herbivores to either avoid or disarm the protective response of plants is one aspect that likely drives this coevolutionary dynamic (Futuyma & Agrawal, 2009b). More recently, herbivorous insects (and other insects) have faced an onslaught of strong selective pressure imposed by synthetic pesticides (Sánchez-Bayo, 2021; Umetsu & Shirai, 2020). Plant defensive chemicals and synthetic insecticides impact the evolution of insect herbivores separately and through their interaction. Understanding these impacts can shed light onto both ancient and modern forces that shape the adaptive landscape in one of the most diverse group of animals on the planet.

Cytochrome P450 monooxygenases or Cyps are a key protein family underlying detoxification of plant defensive chemicals and insecticides in herbivores (Schuler, 2011; J. G. Scott, 1999). Cyps are ancient, and the genes they are encoded by comprise some of the largest gene families in eukaryotes (Nelson, 2018). Cyps most commonly function as detoxification enzymes that catalyze the first step in removing toxins from the body (Phase I) through the addition of an oxygen molecule, which prepares the compound for excretion (Hodges & Minich, 2015). There have been numerous studies on the impacts of Cyp-mediated resistance to insecticides, especially in *Drosophila* species that are often accompanied by transposable element insertions in or near *cyp6g1*. This gene encodes the Cypg1 enzyme, which has broad xenobiotic detoxification capability (Daborn et al., 2002; Chung et al., 2007; Harrop et al., 2014; X. Li et al., 2007; Schlenke & Begun, 2004). Because transposable elements can act as catalysts for increased *cis* or *trans* regulatory expression, it is hypothesized that their insertion enables rapid adaptation to the more recent strong selective pressures of synthetic insecticides (X. Li et al., 2007). However, interestingly, some aspects of these adaptations, namely gene duplication events, predate the onset of widespread synthetic pesticide use (Schmidt et al., 2010).

The preadaption hypothesis posits that selection by plant defensive chemicals may provide herbivores, or insects that interact with plants in other ways (i.e. feeding on microbes within decaying leaves or fruit), with an advantage when presented with novel synthetic insecticides (Gordon, 1961). Comparative tests have concluded that plant diet does influence the ability of insects to detoxify and the advantage gained is likely dependent on the types of plant chemicals with which the insect species is interacting. Insect interactions with plants in the Brassicaceae yielding the highest advantage against insecticides (Hardy et al., 2018). In mustard plants, aliphatic and indolic glucosinolates are the two main precursors of defensive chemicals. Both of these are activated by myrosinases. Aliphatic glucosinolates are transformed into isothiocyanates or mustard oils and are efficiently detoxified by glutathione *S*-transferases (Gloss et al., 2014). Indolic glucosinolates are the precursors of plant defensive chemicals like indole 3-carbinol and possibly cyanide (Dixit et al., 2022).

Cyps underlie detoxification of plant defensive chemicals in many arthropods, including lepidopterans, cactophilic *Drosophila*, and spider mites, among others (Calla et al., 2017; Frank & Fogleman, 1992; Salehipourshirazi et al., 2021; Schuler, 2011). In

some cases, these proteins also confer resistance to defensive compounds created by the indolic glucosinolate pathway. Much of what we know about insect Cyps comes from studies on lepidopterans. In these insects, Cyps are implicated in their ability to detoxify their own host plant chemicals as well as a variety of other chemicals including indole-3-carbinol (X. Li et al., 2002; Stevens et al., 2000). There is evidence that *cyp* duplications can lead to either specialization of the encoded protein to detoxify particular toxins or the evolution of substrate generalism in different copies (X. Li et al., 2004). Generalist Cyps are sometimes able to detoxify indole-3-carbinol even though glucosinolate-derived defensive compounds are not commonly encountered by these proteins. In an experimental evolution study, spider mites evolved the ability to utilize indole-containing plants, and *cyps* were implicated in this adaptation (Salehipourshirazi et al., 2021).

Here we explored the preadaptation hypothesis in a comparative context in *Scaptomyza* and close relatives. *Scaptomyza* is a paraphyletic clade nested within the *Drosophila* that contains several herbivorous lineages. It is estimated that herbivory evolved in this group approximately 10-15 mya (Lapoint et al., 2013; Whiteman et al., 2012). We focused on species that span both herbivorous and non-herbivorous trophic levels. The pre-adaptation hypothesis is relatively untested, and the relatively recent shift to herbivory and well characterized adaptations to insecticides in the clade make *Scaptomyza* an ideal group to utilize to untangle the various threads that can lead to pesticide resistance in herbivorous insects. We first characterized the evolutionary history of *cyp6g1* genes within and among *Scaptomyza* and *Drosophila.* We included *S. flava* and *S.montana,* both mustard specialists, *S. graminum,* which eats plants in the carnation family*, S. pallida* and *S. hsui,* which both eat decaying plants, *D. mojavensis*, which feeds on decaying cactus tissue, and *D. melanogaster, D. virilis and D. grimshawii,* which all feed on decaying fruit or sap fluxes. As larvae, these species all utilize plant tissues, whether fruit or leaves, as habitat, but they vary in whether their main source of nutrients is provided by living leaves or yeast in decaying leaves or fruit. We also laid the groundwork for future research exploring the function of *cyp6g1* copies across this clade and the impact of the P-transposon insertion in the genomic region harboring seven *cyp6g1* copies in *S. flava.*

Methods

Determining cyp6g1 *synteny*

To understand the relative locations and sizes of the tandem *cyp6g1* repeats in *Scaptomyza* (B. Y. Kim et al., 2021; Peláez et al., 2023) and *Drosophila* (Drosophila 12 Genomes Consortium et al., 2007) species, we first annotated *cyp6g1* locations using gene coordinates and labels reported in (Peláez et al., 2023). Briefly, MAKER2 was used to annotate protein coding genes in *S. flava* and predicted *Drosophila* gene sequences and an *S. flava* transcriptome was used to inform the gene models. Genes were subsequently manually curated using iterative BLAST searches. *Cyp6g1* copies are represented by green bars and small gray bars represent pseudogenic *cyp6g1* regions. To identify genes flanking and between *cyp6g1* copies, we performed BLASTn searches of those regions (Johnson et al., 2008). We used an identity match cut off of

99% and e-value of 0.00. When we found genes that overlapped, we depicted the larger of the two, and the omitted genes were also consistent for the species represented here. Sizes of genes and intergenic regions were converted from nucleotide positions to points and depicted in Adobe Illustrator. Species tree topology reflects that reported in (Matsunaga et al., 2022).

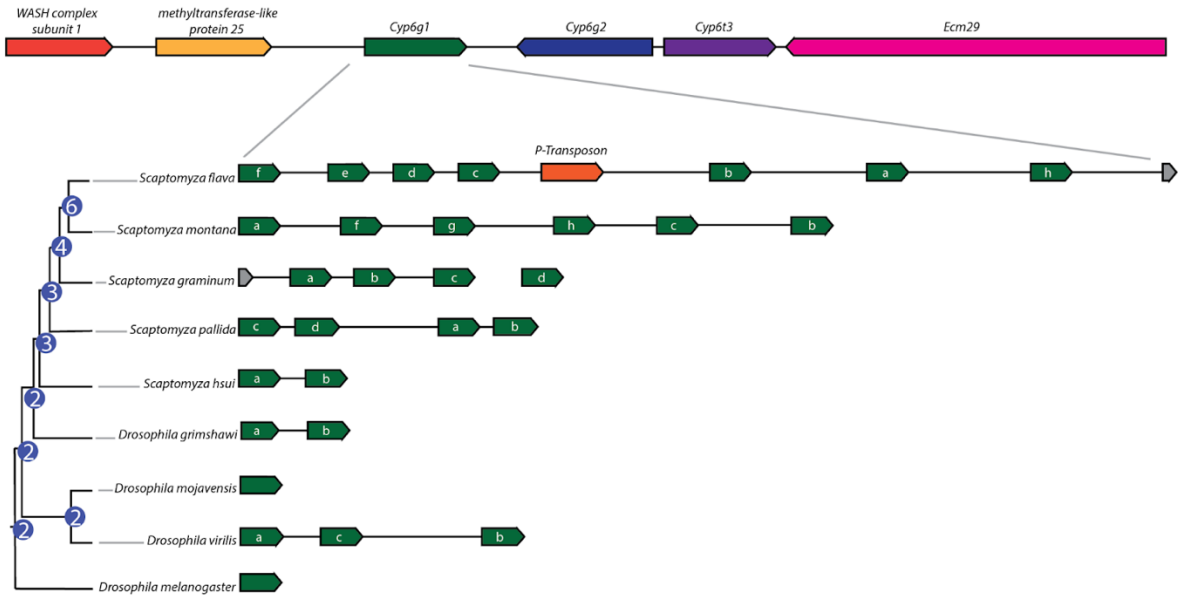


Figure 1 : *cyp6g1* copy number increases in herbivorous lineages. Synteny plots depict physical locations and relative sizes of tandem cyp6g1 copies in select *Scaptomyza* (B. Y. Kim et al., 2021; Peláez et al., 2023) and *Drosophila (Drosophila 12 Genomes Consortium et al., 2007)* species using gene coordinates reported in (Peláez et al., 2023)). All copies are located on the same scaffold except *S. graminum* cyp6g1d, which is found on a different scaffold than the other copies in theis species. Small grey genes represent pseudogenic regions. Numbers in blue circles represent the copy number ancestral state reconstruction as calculated in CAFE (De Bie et al., 2006). Species tree topology reflects that reported in (Matsunaga et al., 2022).

Copy number ancestral state reconstruction

We calculated the number of *cyp6g1* copies in the ancestral state across nodes in the tree using CAFE v5.1.0 (Chen et al., 2018) using default parameters and tree topology and a matrix of *cyp6g1* copy numbers for each species published in (B. Y. Kim et al., 2021; Peláez et al., 2023) (Appendix 3: Supplementary Table 1). We viewed the resulting Base_asr.tre file in iTOL v5 (S. Chen et al., 2018).

Cyp6g1 *gene tree*

For better insight into the phylogenetic relationships between *cyp6g1* copies in our species of interest, we created a maximum likelihood gene tree in IQ-TREE v2.3.0. First, we aligned coding sequences of *cyp6g1* reported in Pelaez et al 2023 using the -codon setting in PRANK (Chen et al., 2018). We then used the NT2AA and NONREV

options to create an amino acid aware ML tree in IQ-TREE with 1000 bootstrap replicates (Hoang et al., 2018; Minh et al., 2020). The random seed number assigned was 857556. We then performed an Approximately Unbiased tree topology test to confirm the root placement calculated originally (S. Chen et al., 2018). To aid in visualization, we subsequently re-rooted the tree at the midpoint in iTOL v5 (Letunic & Bork, 2021).

Divergence-based tests for positive selection

To test whether branches in the *cyp6g1* gene tree have experienced positive selection we used the aBSREL (adaptive Branch-Site Random Effects Likelihood) model in Hyphy v.2.5.60 on the codon-aware aligned *cyp6g1* sequences and gene tree described above (Kosakovsky Pond et al., 2019; Smith et al., 2015) using default parameters. aBSREL calculates the ω rate ratio that is allowed to vary for each branch and each codon site. Although we hypothesize that herbivory has played a role in shaping the evolutionary history of *cyp6g1*, there are also many lines of evidence that positive selection has occurred at *cyp6g1* in non-herbivorous lineages such as *D. melanogaster*, so we tested all branches for positive selection and report Holm-Bonferroni corrected p-values (Appendix 3: Supplementary table 1).

Presence of P-transposon in the cyp6g1 *region in a population of* S. flava

To better understand whether the P-transposon we found inserted into the intergenic region between *S. flava cyp6g1* copies in the reference genome was also present in natural populations, we explored alignments of the region containing *cyp6g1* in a pooled samples of larvae collected from populations of S. flava in Massischusets and New Hampshire in 2013 and 2014. These data was first reported in (Peláez et al., 2022). We first trimmed sequences in fastp with a minimum length requirement of 40 base pairs and minimum quality of 20 (S. Chen et al., 2018). Then, we aligned the sequences using bwa v0.7.17 mem command with default parameters (H. Li, 2013) to the reference genome of *S. flava* (Peláez et al., 2023). After aligning, we generated sorted sam and bam files with bwa sampe and samtools v1.8 view -b and sort, respectively. We manually inspected the alignments to find fragments that map across the gene boundary of the P-transposon.

S. flava cyp6g1 *gene expression*

To determine whether duplicated *cyp6g1* copies in *S. flava* are exhibiting differential expression levels in response to plant defensive chemistry or the inserted transposon between *cyp6g1c* and *cyp6g1b*, we compared expression levels of each gene in whole body and gut samples using data from Whiteman et al. (2012) and Gloss et al. (2019), respectively. Larvae for whole body samples were reared for 5 days on two *Arabidopsis thaliana* genotypes: 1) wild type glucosinolate-containing plants (Col-0) and 2) mutant plants with four glucosinolate genes knocked out (GKO; *myb28myb29cyp79b2cyp79b3*). These genes encode for the aliphatic and indolic

glucosinolate pathway, but the aromatic (benzenic) glucosinolate pathway remain intact. We accessed these Illumina sequencing data from NCBI accession number SRR514175. Larvae for gut samples were reared for 3 or 12 hours on Pad-3 or GKO. Pad-3 plants which have the pathways for aliphatic and indolic glucosinolates intact, but not a sidebrach of the indolic glucosinolate pathway that produces camalexin. Camalexin is an antimicrobial compound that is also eliminated in GKO plants, since the entire tryptophan-derived, indolic glucosinolate pathway is knocked out. After rearing on these plants with or without glucosinolates, midguts were dissected out and flash frozen before RNA was extracted. We accessed these Illumina sequencing data from NCBI accession numbers SRR7649192-SRR7649195.

We first separated the fastq file containing whole-body sequences into experimental replicates using grep. We trimmed sequences using default settings in fastp (S. Chen et al., 2018). We then aligned the files in hisat2 v.2.1.0 to the indexed *S. flava* reference genome assembly (GCA_030179655.1; (D. Kim et al., 2019). We quantified expression levels with htseq v0.9.1 using the htseq-count command (Putri et al., 2022). We analyzed the resulting count data in DESeq2 v. 1.42.1 using the design ~ batch + condition, with conditions being the plant chemistry type (Love et al., 2014). We then used the heatmap function in base R v2023.12.0+369 (R Core Team, 2021) to visualize the data.

Metabolism of insecticides by D. melanogaster cyp6g1

We had the *cyp6g1* coding sequences from *S. flava, S. graminum, S. hsui, and D melanogaster* published in Pelaez et al, 2023 cloned into pENTR by Twist Biosciences. We transformed this plasmid into the XL1 Blue strain of *E. coli*, which was made competent using the Zymo *Mix & Go! E. coli* Transformation Kit and Buffer Set. We performed the Gateway™ LR Clonase™ II reaction to catalyze recombination between our entry pENTR plasmid and the destination pAWF plasmid, which contains a 3' FLAG tag (Figure 4a). We then transfected the purified plasmid into *Drosophila* S2 cells acquired from the UC Berkeley Cell Culture Facility using Cellfectin II and the associated protocol. Briefly, we placed ~8e5 cells in 2.5 ml of plain, unsupplemented ESF 921 Insect Culture Broth (no antibiotics) in 6 well plate. We mixed 8 µl of Cellfectin II into 100 µl of unsupplemented broth and incubated at room temperature for 5 minutes. We also mixed 3 µg of purified plasmid DNA in 100 µl of unsupplemented broth and incubated at room temperature for 5 minutes. We then combined the two mixtures and incubated at room temperature for 15 minutes. After incubation, we added the mixture dropwise onto the cells and gently swirled. We incubated the transfection mixture for 3 hours at 27 degrees C. We then removed the transfection media and replaced it with 2.5 ml of supplemented broth containing antibiotics and incubated for another 21-22 hours. We then added 400 ppm of dinotefuran (Thermo Scientific) to transfected cells and non-transfected control cells and incubated for 48 hours. At that time, we transferred to microcentrifuge tubes, spun cells down, and removed the supernatant. We measured the amount of dinotefuran left in the supernatant in triplicate on an LC-MS machine. We also validated the expression of cyp6g1 with RT-PCR and a

western blot (Supplemental Figure 1). We compared the dinotefuran amounts remaining in the wells with Welch's two sample t-test in R (R Core Team, 2021).

Results

Cyp6g1 *copy number ancestral state reconstruction*

We ran CAFE v5.1.0 to estimate the ancestral copy numbers in our species of interest (Figure 1). We found expansions of *cyp6g1* along most lineages, except for the branch leading to *S. hsui,* which lost one copy of *cyp6g1*. There was one duplication event leading to the non-herbivorous *Scaptomyza*, and subsequent duplications at each node between non-herbivorous *Scaptomyza* and *S. flava*, which has the largest number of extant copies (7, plus one pseudogenic copy). We found the cumulative rate of gains and losses per gene per unit time (λ) to be 0.011 for this gene family. This represents a large increase over the rate of turnover for all detoxification genes in herbivores (λ = 0.005) reported in Peláez et al., 2023. In this study, the authors also validated that these increases are not caused by the long branch length leading to the herbivorous clade, and so a long branch is also unlikely to play a role here.

Presence of P-transposon in the cyp6g1 *region in a population of* S. flava

We were interested in understanding whether the P-transposon that is present in the reference genome between *cyp6g1b* and *cyp6g1c* may have been swept through the population, perhaps indicating strong, recent selective pressure, like that imposed by synthetic insecticide use. As an initial pass to determine the distribution of this gene through time and space, we mapped 3 separate pooled sequencing data sets to the *S. flava* genome and looked for fragments that spanned the gene boundary. We found the P-transposon insertion in three of three of the data sets (Table 1). Further study is necessary to determine the wider distribution of this insertion, and the underpinning of its persistence.

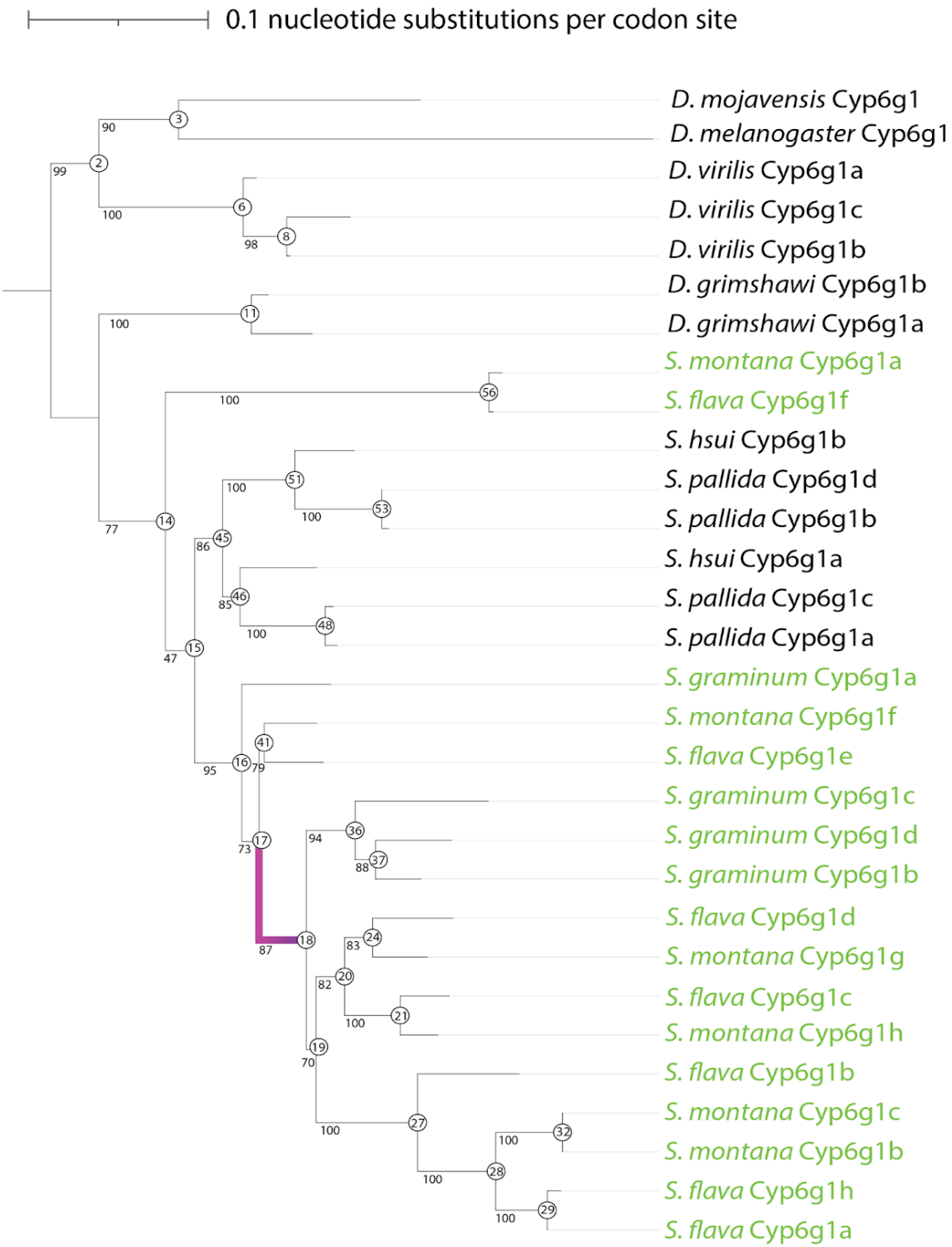


Figure 2 : ML *cyp6g1* tree with evidence for positive selection among herbivorous copies. gene tree depicting the phylogenetic relationships between *cyp6g1* copies in select Drosophila and Scaptomyza species. Created from a condon aware alignment in IQ-TREE v 2.5.60 with NONREV substitution model. Bootstrap support is depicted along the branches. Gene copies in herbivorous flies are notated in green font. The branch undergoing positive selection according to the Hyphy aBSREL model is highlighted in magenta. Node labels from Hyphy are noted in black.

cyp6g1 *gene tree*

To determine the evolutionary relationships between cyp6g1 copies in select herbivorous and non-herbivorous drosophilid species, we created an amino acid aware gene tree for *cyp6g1* in IQ-TREE (Figure 2). The topology of the tree generally mirrors the topology of the species tree with variance in the topology of the *Scaptomyza* branches where the copy number increases rapidly. The rooting position of the tree calculated by IQ-TREE places sequences in the *Drosophila* as sister to one another, which does not reflect the species tree. The log likelihood of the tree was -7122.65 (s.e. 271.89) compared to the unconstrained log-likelihood of 3054.79. The number of free parameters (# of branches + # of model parameters) was 437.

Divergence-based tests for positive selection

We hypothesized that possible neofunctionalization of some copies of *cyp6g1* in herbivorous species would require positive selection of that branch. To better understand whether particular branches in the tree were undergoing selection, we used the aBSREL model in Hyphy to calculate the branch-site values of ω for all of the branches in the *cyp6g1* gene tree. Since we performed tests on all branches, we report Holm-Bonferroni corrected p-values. After this correction, we found that one branch of the tree is undergoing positive selection. This branch leads to most of the *cy6g1* copies in the herbivorous species included in the tree (Figure 2). Other notable, but ultimately non-significant branches include Node 24, leading to *S. flava cyp6g1d* and *S. montana cyp6g1g*, and the *S. melanogaster cyp6g1* branch (Appendix 3: Supplementary Table 1).

S. flava cyp6g1 *gene expression*

To better understand the impact of plant defensive chemicals and the transposon inserted between *cyp6g1c* and *cyp6g1b* copies on expression levels of these genes, we compared $\log_2$ fold change in expression levels in midgut and whole body tissues between the seven *cyp6g1* duplicates in response to the presence of glucosinolates (Figure 3). We produced alignments that varied between ~9-30 million reads depending on the experimental replicate (Appendix 3: Supplementary Table 2). For both experiments, larvae were reared on GKO plants and transplanted to either plants with glucosinolates or another GKO plant. Gut analysis was completed on larvae that ate their experimental plant for 3 or 12 hours while whole body analysis was completed on larvae that ate their experimental plant for 5 days. This difference in experimental design may underlie the differences in significance of *cyp6g1* expression we found across tissue types. Through a DESeq2 analysis, we found one *cyp6g1* copy (*cyp6g1h*) had significantly increased gene expression in whole body samples in response to the presence of glucosinolates in the diet with a $\log_2$ fold increase of 2.01 ($p=0.01$) relative to expression in larvae reared on plants without glucosinolates. Other *cyp6g1* copies either were not found in our data sets (*cyp6g1a*) or did not have significant changes in

expression when comparing between larvae that encountered plant toxins and larvae that did not.

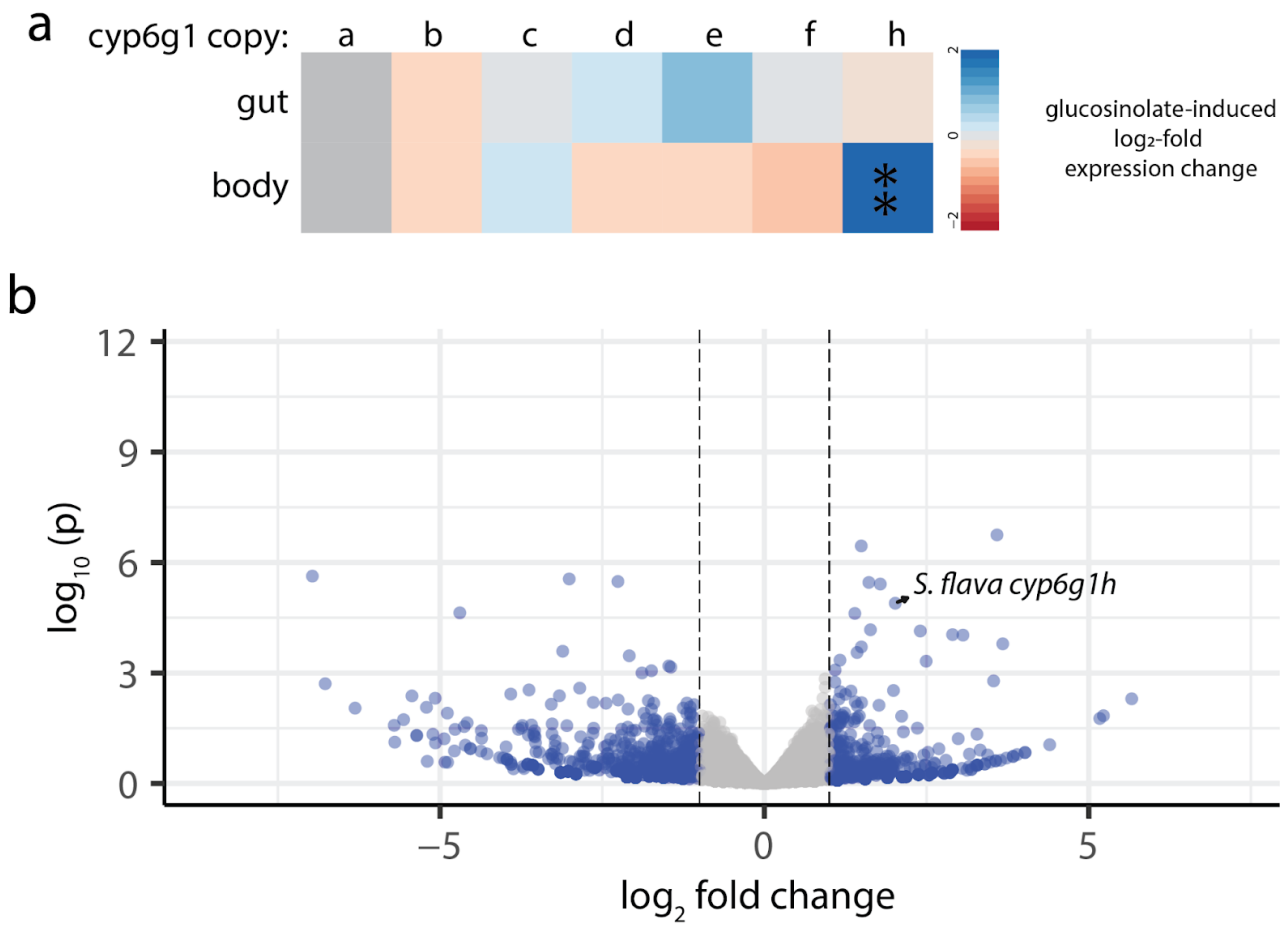


Figure 3: S. flava *cyp6g1h* has increased expression in response to plant defensive chemicals in a DESeq analysis a) Heat map depicts the log2 fold expression level changes in larval midgut and whole body tissues in response to plant chemicals (either with or without glucosinolates). Color scale range is set to depict the largest expression change measured in the *cyp6g1* copies we focused on. b) Volcano plot of whole body expression data with *cyp6g1h* labeled. Blue dots represent all up or down regulated genes.

Metabolism of insecticides by D. melanogaster cyp6g1

To test whether the transfected copy of *cyp6g1* was able to detoxify the pesticide dinotefuran, we compared the amount of dinotefuran in wells with transfected *cyp6g1* and without. We found differences in the amount of dinotefuran remaining (Figure 4b; t = 27.953, df = 5.4043, p = $4.695e^{-07}$). In future experiments, we will test whether *cyp6g1* copies in other lineages (*S. flava, S. graminum, S. hsui*) can detoxify pesticides and/or plant defensive chemicals derived from indolic glucosinolates.

Location	Collection date	SRA Accession	P-transposon present?
Belmont, MA	June & July 2013	SRR15275350-SRR15275365	Yes
Belmont, MA	June & July 2014	SRR20722523 & SRR20722528	Yes
Portsmouth, NH	June & July 2014	SRR20722524 & SRR20722525	Yes

Table 1: P-transposon insertion is present in 3 *S. flava* populations in New England

Discussion

The evolution of herbivory requires multimodal and complex adaptations due to the dynamic nature of coevolving plant and animal partners. Interactions between plants and their herbivores shape the adaptive landscape of both lineages (Southwood, 1972). A major source of this exchange is through the evolution of plant defensive chemicals and the subsequent counter-evolution of adaptations by herbivores (Ehrlich & Raven, 1964). Adaptations in detoxification repertoires can also confer unexpected advantages, like the ability to detoxify other xenobiotic chemicals in the environment. This unexpected advantage is the basis of the pre-adaptation hypothesis, which posits that selection by plant defensive chemicals pre-adapts insects subsequently presented with insecticides (Gordon, 1961). General patterns indicate that insects that attack plants in the Brassicaceae are more likely to evolve insecticide resistance (Hardy et al., 2018).

We focused our efforts on Cytochrome P450 proteins, which have been shown to be involved in detoxification of both plant and insecticide xenobiotics in arthropods. The defensive chemicals in Brassicaceae are diverse and vary from plant to plant, but they can be divided into classes depending on their amino acid precursor. Aliphatic and indolic glucosinolates are among the most common. Upon plant damage, both of these non-toxic chemicals are activated by myrosinases into a variety of toxic, defensive, and

plant-protecting chemicals. Aliphatic glucosinolate produced isothiocyanates are efficiently detoxified by glutathione *S*-transferases, but also activate *cyp6g1* transcription in some contexts (Gloss et al., 2014; Vázquez-Gómez et al., 2010). Indolic glucosinolate products like indole-3-carbinol interact more readily with Cyps. Cyps with generalist substrate binding in lepidopterans showed affinity for indole-3-carbinol (X. Li et al., 2004, 2002; Stevens et al., 2000). Additionally, adaptation in glucosinolate-naïve spider mite *cyps* confered the ability to utilize indole-containing plants as hosts (Salehipourshirazi et al., 2021).

Here, we focused on *cyp6g1*, which underlies major adaptations to insecticides in *Drosophila* species. In vitro studies have validated that Cp6g1 proteins can detoxify a variety of insecticides (Joussen et al., 2008; Sparks et al., 2012). Transposable element insertions upstream of *cyp6g1* increase expression levels in *D. melanogaster* and *D. simulans*, which has been implicated in DDT resistance with further studies indicating cross-resistance to other insecticides (P. Daborn et al., 2001; P. J. Daborn et al., 2002; Schlenke & Begun, 2004). In *D. melanogaster*, expression of *cyp6g1* is upregulated in a tissue-specific manner, with tissues in the midgut, Malpighian tubules, and fat body mostly affected (Chung et al., 2007). In addition to expression level changes, duplications of *cyp6g1* have also been linked to insecticide resistance (Harrop et al., 2014; Schmidt et al., 2010).

Cyp6g1 has undergone multiple rounds of duplication in closely related herbivorous lineages in the genus *Scaptomyza* (Figure 1). CAFE analysis places duplication events at the ancestor leading to the herbivorous lineage, which split approximately 10-15 mya (Figure 1; (Lapoint et al., 2013; Whiteman et al., 2012)). Duplication events have contributed to other rapid adaptive events in *D. melanogaster*, but like the *Scaptomyza* species studied here, these duplications occurred long before the advent of widespread insecticide use indicating there was a selective pressure shaping these detoxification genes on a longer timescale in addition to the very strong selective pressure imposed by insecticides (Schmidt et al., 2010).

One approach to estimate the adaptive landscape experienced by candidate genes, including *cyp6g1,* is through a phylogenetic and divergence-based approach (Figure 2). Accordingly, we used branch-site tests that calculate the ratio of synonymous to non-synonymous substitutions or ω to test whether *cyp6g1* copies were undergoing positive selection. We are most interested in whether herbivory has spurred adaptation in this gene repertoire, but although the power to detect positive selection is reduced, we performed these tests without *a priori* designation of test branches because of the overwhelming evidence that other, non-herbivorous *Drosophila cyp6g1* genes have also undergone strong selective events. Despite this, we found that the branch leading to a majority of herbivore copies was likely under positive selection. Like *S. flava* and *S. montana*, *S. graminum* is also herbivorous, but is specialized on plants in the Caryophyllaceae, which have a different defensive chemical profile than the mustards upon which *S. flava* and *S. montana* feed. This finding suggests that herbivory in itself imposed strong selection pressure on *cyp6g1* pressure, irrespective of the particular plant defensive chemistry involved.

We observed an O-type P-transposon between *cyp6g1c* and *cyp6g1b* in the reference genome of *S. flava.* Although, like most new mutations, transposable element insertions generally lower fitness of flies, some are beneficial with respect to fitness and have often been implicated in the evolution of xenobiotic resistance in insects (X. Li et al., 2007; Mackay, 1989; Mérel et al., 2020; Schlenke & Begun, 2004). To better understand whether this P-transposon is present across time and geographical space, we searched for its presence in pooled sequencing data taken one year apart at the same location and at the same time, but different locations. We found the transposon in all three of the data sets. More investigation is necessary to understand whether this insertion is fixed in these and other populations and whether this site is involved in the detoxification of xenobiotics in *S. flava*.

One hypothesis is that these transposable elements may work in *cis* to increase transcription of their neighboring gene (Schmidt et al., 2010). To address this hypothesis in a crude way, we measured the change in expression of these genes in larvae feeding on *A. thaliana* plants with or without plant defensive chemicals known as glucosinolates. We found that one paralogous copy of *cyp6g1* (*cyp6g1h*) showed increased expression in response to plant chemicals in whole body tissue samples (Figure 3). Although we hypothesize that plant chemistry has shaped the evolution of the *cyp6g1* copies in herbivores, it is possible that selection in the P-transposon has occurred more recently and in response to other selective pressures (i.e. insecticides), especially in light of the many instances of involvement in insecticide resistance in close relatives. The absence of this insertion in the closely related *S. montana* and the absence of changes in expression in response to plant chemistry in neighboring *cyp6g1* copies indicate that plant defensive chemicals are likely not driving selection in the P-transposon.

Further study is required to identify the impact of insecticides on expression in other tissues, life stages and contexts, as well as the presence of the P-transposon in other populations of *S. flava* and its possible impact on insecticide resistance. We plan to test the efficacy of *S. flava* Cyp6g1 proteins at detoxifying both plant chemicals and insecticides *in vitro* using proteins heterologously expressed in Drosophila S2 cells. Using the assay described here, we will be able to begin to untangle the effects of plant defensive chemicals and insecticides on the evolution of *cyp6g1* in *Scaptomyza*.

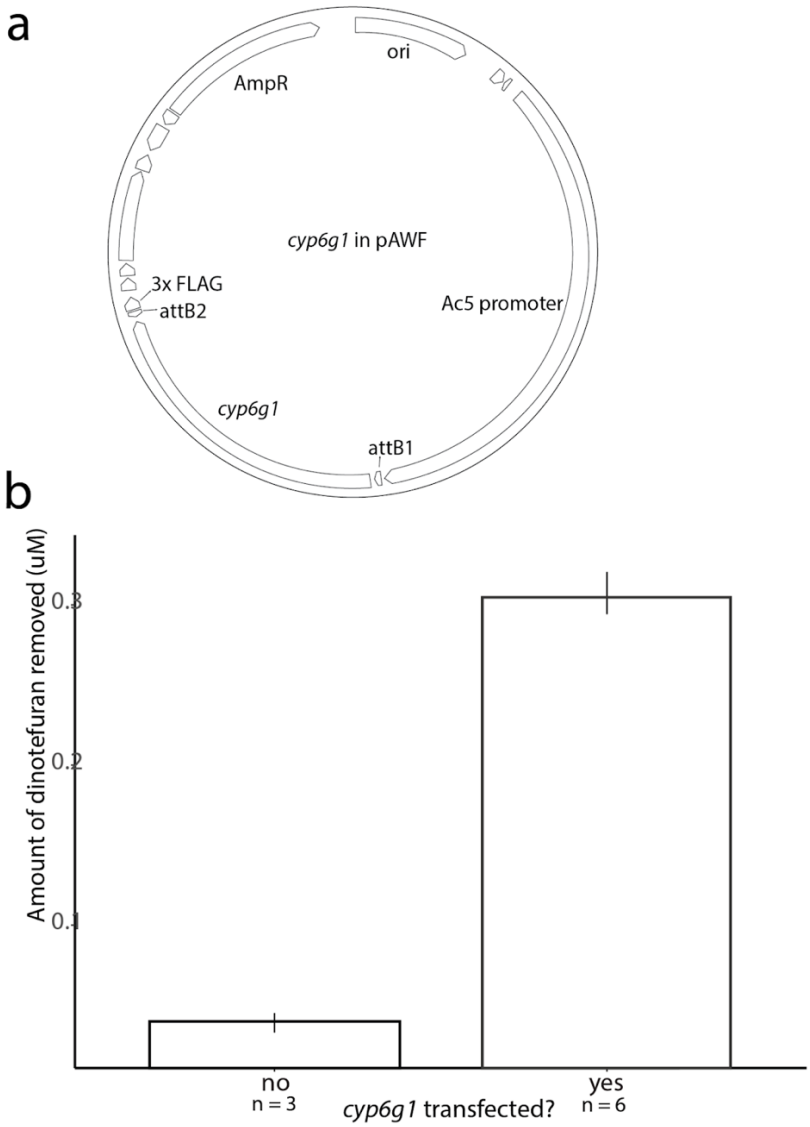


Figure 4 : *D. melanogaser cyp6g1* detoxifies dinotefuran a) map of plasmid used to transfect S2 cells with a copy of *cyp6g1*. **b).** We measured the amount of dinotefuran removed by transfected S2 cells in culture.

CONCLUSION

Throughout my dissertation, I have focused on the evolutionary leaps required by insect herbivores to use plants as food and habitat. Generally, herbivory requires adaptive shifts in behavioral, morphological, nutritional and detoxification traits to overcome the defenses that plants present. In all of my chapters, I focused on understanding these adaptations in *Scaptomyza flava,* which is a leaf-mining, drosophilid fly, In Chapter 1, I focused on understanding behavioral, morphological, and nutritional shifts required for herbivory, and in Chapters 2 and 3, I focused on understanding adaptations that enable detoxification of plant and other toxic chemicals that herbivores encounter. Chapters 1 and 3 both utilized the powerful comparative framework presented by the trophic niches of these flies and their relatives–from completely non-herbivorous to obligately so–to understand how traits vary along a range of plant interactions. Chapter 2 focused on the multifaceted role of the microbiome in herbivore plant interactions. Overall, this work has contributed “insights into the evolution of herbivory from a leaf-mining fly.”

References – Chapter 1

Aguilar, J. M. 2022, December 6. jmagui/HerbivoryPaper: Sflava_HerbivoryPaper.

Behmer, S. T. 2017. Overturning dogma: tolerance of insects to mixed-sterol diets is not universal. Current opinion in insect science 23:89–95.

Behmer, S. T., and W. David Nes. 2003. Insect Sterol Nutrition and Physiology: A Global Overview. Pages 1–72 Advances in Insect Physiology. Academic Press.

Bernays, E. A., and R. E. Chapman. 1994. Host-Plant Selection by Phytophagous Insects. Springer, Boston, MA.

Bing, X., J. Gerlach, G. Loeb, and N. Buchon. 2018. Nutrient-Dependent Impact of Microbes on Drosophila suzukii Development. mBio 9.

Chandler, J. A., J. A. Eisen, and A. Kopp. 2012. Yeast communities of diverse Drosophila species: comparison of two symbiont groups in the same hosts. Appl. Environ. Microbiol. 78: 7327–7336.

Clark, A. J., and K. Block. 1959. The absence of sterol synthesis in insects. The Journal of biological chemistry 234:2578–2582.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and Plants: A Study in Coevolution. Evolution; international journal of organic evolution 18:586–608.

Fagan, W. F., E. Siemann, C. Mitter, R. F. Denno, A. F. Huberty, H. A. Woods, and J. J. Elser. 2002. Nitrogen in insects: implications for trophic complexity and species diversification. The American naturalist 160:784–802.

Feder, J. L., C. A. Chilcote, and G. L. Bush. 1988. Genetic differentiation between sympatric host races of the apple maggot fly Rhagoletis pomonella. Nature 336:61–64.

Feng, S., S. Liu, Z. Luo, and K. Tang. 2015. Direct saponification preparation and analysis of free and conjugated phytosterols in sugarcane (Saccharum officinarum L.) by reversed-phase high-performance liquid chromatography. Food chemistry 181:9–14.

German, D. P., and M. H. Horn. 2006. Gut length and mass in herbivorous and carnivorous prickleback fishes (Teleostei: Stichaeidae): ontogenetic, dietary, and phylogenetic effects. Marine biology 148:1123–1134.

Gloss, A. D., D. G. Vassão, A. L. Hailey, A. C. Nelson Dittrich, K. Schramm, M. Reichelt, T. J. Rast, A. Weichsel, M. G. Cravens, J. Gershenzon, W. R. Montfort, and N. K. Whiteman. 2014. Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the drosophilidae. Molecular biology and evolution 31:2441–2456.

Godoy-Herrera, R., B. Burnet, K. Connolly, and J. Gogarty. 1984. The development of locomotor activity in Drosophila melanogaser larvae. Heredity 52:63–75.

Goldman-Huertas, B., R. F. Mitchell, R. T. Lapoint, C. P. Faucher, J. G. Hildebrand, and N. K. Whiteman. 2015. Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. Proceedings of the National Academy of Sciences of the United States of America 112:3026–3031.

Gompert, Z., L. K. Lucas, C. C. Nice, J. A. Fordyce, M. L. Forister, and C. A. Buerkle. 2012. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. Evolution; international journal of organic evolution 66:2167–2181.

Griffen, B. D., and H. Mosblack. 2011. Predicting diet and consumption rate differences between and within species using gut ecomorphology. The Journal of animal ecology 80:854–863.

Groen, S. C., P. T. Humphrey, D. Chevasco, F. M. Ausubel, N. E. Pierce, and N. K. Whiteman. 2016. Pseudomonas syringae enhances herbivory by suppressing the reactive oxygen burst in Arabidopsis. Journal of insect physiology 84:90–102.

Hackman, W. 1959. On the genus Scaptomyza Hardy (Dipt., Drosophilidae) with descriptions of new species from various parts of the world. Acta Zoologica Fennica:3–73.

Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94–98.

Hothorn, T., K. Hornik, M. A. Van De Wiel, A. Zeileis, and Others. 2008. Implementing a class of permutation pests: the coin package.

Ikekawa, N., M. Morisaki, and Y. Fujimoto. 1993. Sterol metabolism in insects: dealkylation of phytosterol to cholesterol. Accounts of chemical research 26:139–146.

Jaenike, J. 1978. On optimal oviposition behavior in phytophagous insects. Theoretical population biology 14:350–356.

Janson, E. M., R. J. Grebenok, S. T. Behmer, and P. Abbot. 2009. Same Host-Plant, Different Sterols: Variation in Sterol Metabolism in an Insect Herbivore Community. Journal of chemical ecology 35:1309.

Jing, X., and S. T. Behmer. 2020. Insect Sterol Nutrition: Physiological Mechanisms, Ecology, and Applications. Annual review of entomology 65:251–271.

Jing, X., R. J. Grebenok, and S. T. Behmer. 2013. Sterol/steroid metabolism and absorption in a generalist and specialist caterpillar: effects of dietary

sterol/steroid structure, mixture and ratio. Insect biochemistry and molecular biology 43:580–587.

Jing, X., H. Vogel, R. J. Grebenok, K. Zhu-Salzman, and S. T. Behmer. 2012. Dietary sterols/steroids and the generalist caterpillar Helicoverpa zea: physiology, biochemistry and midgut gene expression. Insect biochemistry and molecular biology 42:835–845.

Karasov W H, Martínez del Rio C, and Caviedes-Vidal E. Ecological physiology of diet and digestive systems. (2011) PMID: 21314432

Kim, B. Y., J. R. Wang, D. E. Miller, O. Barmina, E. Delaney, A. Thompson, A. A. Comeault, D. Peede, E. R. D'Agostino, J. Pelaez, J. M. Aguilar, D. Haji, T. Matsunaga, E. E. Armstrong, M. Zych, Y. Ogawa, M. Stamenković-Radak, M. Jelić, M. S. Veselinović, M. Tanasković, P. Erić, J.-J. Gao, T. K. Katoh, M. J. Toda, H. Watabe, M. Watada, J. S. Davis, L. C. Moyle, G. Manoli, E. Bertolini, V. Košťál, R. S. Hawley, A. Takahashi, C. D. Jones, D. K. Price, N. Whiteman, A. Kopp, D. R. Matute, and D. A. Petrov. 2021. Highly contiguous assemblies of 101 drosophilid genomes. eLife 10.

Lapoint, R. T., P. M. O'Grady, and N. K. Whiteman. 2013. Diversification and dispersal of the Hawaiian Drosophilidae: the evolution of Scaptomyza. Molecular phylogenetics and evolution 69:95–108.

Li, S., and X. Jing. 2020. Fates of dietary sterols in the insect alimentary canal. Current opinion in insect science 41:106–111.

Lyu, F., X.-X. Hai, Z.-G. Wang, and Y. Bi. 2018. Influence of visual cues on oviposition site searching and learning behavior in the parasitic beetle Dastarcus helophoroides (Fairmaire) (Coleoptera: Bothrideridae). Scientific reports 8:17331.

Martin, A. M. 2015. The effects of sterols on Drosophila melanogaster: physiology and biochemistry. Master of Science, Texas A&M University.

Martin, N. A. 2004. History of an invader, Scaptomyza flava (Fallen, 1823) (Diptera: Drosophilidae). New Zealand journal of zoology 31:27–32.

Martin, N. A. 2012. Ecological studies of Scaptomyza flava (Fallen, 1823) (Diptera: Drosophilidae) and its parasitoids. The New Zealand entomologist 35:58–67.

Matsunaga, T., C. E. Reisenman, B. Goldman-Huertas, P. Brand, K. Miao, H. C. Suzuki, K. I. Verster, S. R. Ramírez, and N. K. Whiteman. 2022. Evolution of Olfactory Receptors Tuned to Mustard Oils in Herbivorous Drosophilidae. Molecular biology and evolution 39.

McCutchan, J. H., W. M. Lewis, C. Kendall, and C. C. McGrath. 2003. Variation in Trophic Shift for Stable Isotope Ratios of Carbon, Nitrogen, and Sulfur. Oikos 102:378–390.

McPheron, B. A., D. C. Smith, and S. H. Berlocher. 1988. Genetic differences between host races of Rhagoletis pomonella. Nature 336:64–66.

Mitchell-Olds, T. 2001. Arabidopsis thaliana and its wild relatives: a model system for ecology and evolution. Trends in ecology & evolution 16:693–700.

Mitter, C., B. Farrell, and B. Wiegmann. 1988. The Phylogenetic Study of Adaptive Zones: Has Phytophagy Promoted Insect Diversification? The American naturalist 132:107–128.

Nagaya, H., F. J. Stewart, and M. Kinoshita. 2021. Swallowtail Butterflies Use Multiple Visual Cues to Select Oviposition Sites. Insects 12.

O'Connor, T. K., P. T. Humphrey, R. T. Lapoint, N. K. Whiteman, and P. M. O'Grady. 2014. Microbial interactions and the ecology and evolution of Hawaiian Drosophilidae. Frontiers in microbiology 5:616.

Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. De Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. Ribeiro Cunha, T. Smith, A. Stier, C. J. F. Ter Braak, and J. Weedon. 2022. vegan: Community Ecology Package.

Peláez, J. N., A. D. Gloss, B. Goldman-Huertas, B. Kim, R. T. Lapoint, G. Pimentel-Solorio, K. I. Verster, J. M. Aguilar, A. C. Nelson Dittrich, M. Singhal, H. C. Suzuki, T. Matsunaga, E. E. Armstrong, J. L. M. Charboneau, S. C. Groen, D. H. Hembry, C. J. Ochoa, T. K. O'Connor, S. Prost, S. Zaaijer, P. D. Nabity, J. Wang, E. Rodas, I. Liang, and N. K. Whiteman. 2023. Evolution of chemosensory and detoxification gene families across herbivorous Drosophilidae. G3 .

Peláez, J. N., A. D. Gloss, J. F. Ray, S. Chaturvedi, D. Haji, J. L. M. Charboneau, K. I. Verster, and N. K. Whiteman. 2022. Evolution and genomic basis of the plant-penetrating ovipositor: a key morphological trait in herbivorous Drosophilidae. Proceedings of the Royal Society B: Biological Sciences 289:20221938.

Perkins, M. J., R. A. McDonald, F. J. F. van Veen, S. D. Kelly, G. Rees, and S. Bearhop. 2014. Application of Nitrogen and Carbon Stable Isotopes (δ15N and δ13C) to Quantify Food Chain Length and Trophic Structure. PloS one 9:e93281.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, O. W., and R. G. Davies. 1977. The Alimentary Canal, Nutrition and Digestion. Pages 192–208 *in* O. W. Richards and R. G. Davies, editors. IMMS' General Textbook of Entomology: Volume I: Structure, Physiology and Development. Springer Netherlands, Dordrecht.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature methods 9:671–675.

Schoonhoven, L., J. J. A. van Loon, and M. Dicke. 2005. Insect-Plant Biology. Oxford University Press.

Soria-Carrasco, V., Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, S. P. Egan, B. J. Crespi, and P. Nosil. 2014. Stick insect genomes reveal natural selection's role in parallel speciation. Science 344:738–742.

Southwood, T. R. E. 1972. The insect/plant relationship--an evolutionary perspective. Royal Entomological Society London Symposium 6:3–30.

Spence, K. O., and J. A. Rosenheim. 2005. Isotopic enrichment in herbivorous insects: a comparative field-based study of variation. Oecologia 146:89–97.

Sterner, R. W., and J. J. Elser. 2002. The Biology of Elements from Molecules to the Biosphere. Princeton University Press.

Thompson, B. M., R. J. Grebenok, S. T. Behmer, and D. S. Gruner. 2013. Microbial symbionts shape the sterol profile of the xylem-feeding woodwasp, Sirex noctilio. Journal of chemical ecology 39:129–139.

Vertacnik, K. L., and C. R. Linnen. 2017. Evolutionary genetics of host shifts in herbivorous insects: insights from the age of genomics. Annals of the New York Academy of Sciences 1389:186–212.

Vervust, B., P. Pafilis, E. D. Valakos, and R. Van Damme. 2010. Anatomical and physiological changes associated with a recent dietary shift in the lizard Podarcis sicula. Physiological and biochemical zoology: PBZ 83:632–642.

Vijendravarma, R. K., S. Narasimha, and T. J. Kawecki. 2013. Predatory cannibalism in Drosophila melanogaster larvae. Nature communications 4:1789.

Whiteman, N. K., A. D. Gloss, T. B. Sackton, S. C. Groen, P. T. Humphrey, R. T. Lapoint, I. E. Sønderby, B. A. Halkier, C. Kocks, F. M. Ausubel, and N. E. Pierce. 2012. Genes involved in the evolution of herbivory by a leaf-mining, Drosophilid fly. Genome biology and evolution 4:900–916.

Whiteman, N. K., S. C. Groen, D. Chevasco, A. Bear, N. Beckwith, T. R. Gregory, C. Denoux, N. Mammarella, F. M. Ausubel, and N. E. Pierce. 2011. Mining the plant-herbivore interface with a leafmining Drosophila of Arabidopsis. Molecular ecology 20:995–1014.

Wilson, A. C. C., L. da S. L. Sternberg, and K. B. Hurley. 2011. Aphids alter host-plant nitrogen isotope fractionation. Proceedings of the National Academy of Sciences of the United States of America 108:10220–10224.

Wipfler, B., K. Schneeberg, A. Löffler, F. Hünefeld, R. Meier, and R. G. Beutel. 2013. The skeletomuscular system of the larva of Drosophila melanogaster (Drosophilidae, Diptera): a contribution to the morphology of a model organism. Arthropod structure & development 42:47–68.

Yang, Y., and A. Joern. 1994. Gut Size Changes in Relation to Variable food Quality and Body Size in Grasshoppers. Functional ecology 8:36–45.

References Chapter 2

Aguilar, J. M., Gloss, A. D., Suzuki, H. C., Verster, K. I., Singhal, M., Hoff, J., Grebenok, R., Nabity, P. D., Behmer, S. T., & Whiteman, N. K. (2024). Insights into the evolution of herbivory from a leaf-mining, drosophilid fly. *Ecosphere*.

Apprill, A., Weber, L. G., & Santoro, A. E. (2016). Distinguishing between Microbial Habitats Unravels Ecological Complexity in Coral Microbiomes. *mSystems*, *1*(5). https://doi.org/10.1128/mSystems.00143-16

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *19*(5), 455–477.

Baumann, P., Lai, C., Baumann, L., Rouhbakhsh, D., Moran, N. A., & Clark, M. A. (1995). Mutualistic associations of aphids and prokaryotes: biology of the genus buchnera. *Applied and Environmental Microbiology*, *61*(1), 1–7.

Berenbaum, M. R. (1988). Allelochemicals in insect-microbe-plant interactions; agents provocateurs in the coevolutionary arms race. In P. B. Deborah K. Letourneau (Ed.), *Novel Aspects of Insect-Plant Interactions* (pp. 97–124). Wiley.

Berg, M., Zhou, X. Y., & Shapira, M. (2016). Host-Specific Functional Significance of Caenorhabditis Gut Commensals. *Frontiers in Microbiology*, *7*, 1622.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Titus Brown, C., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Gregory Caporaso, J. (2018). *QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science* (No. e27295v2). PeerJ Preprints. https://doi.org/10.7287/peerj.preprints.27295v2

Bones, A. M., & Rossiter, J. T. (1996). The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiologia Plantarum*, *97*(1), 194–208.

Bost, A., Martinson, V. G., Franzenburg, S., Adair, K. L., Albasi, A., Wells, M. T., & Douglas, A. E. (2018). Functional variation in the gut microbiome of wild Drosophila populations. *Molecular Ecology*, *27*(13), 2834–2845.

Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* , *33*(19), 3137–3139.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583.

Ceja-Navarro, J. A., Vega, F. E., Karaoz, U., Hao, Z., Jenkins, S., Lim, H. C., Kosina, P., Infante, F., Northen, T. R., & Brodie, E. L. (2015). Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. *Nature Communications*, *6*(1), 1–9.

Chen, J., Ullah, C., Reichelt, M., Beran, F., Yang, Z.-L., Gershenzon, J., Hammerbacher, A., & Vassão, D. G. (2020). The phytopathogenic fungus Sclerotinia sclerotiorum detoxifies plant glucosinolate hydrolysis products via an isothiocyanate hydrolase. *Nature Communications*, *11*(1), 3090.

Cline, J. D. (1969). SPECTROPHOTOMETRIC DETERMINATION OF HYDROGEN SULFIDE IN NATURAL WATERS. *Limnology and Oceanography*, *14*, 454–458.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* , *27*(8), 1164–1165.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772.

Davis, N. M., Proctor, D., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*. https://doi.org/10.1101/221499

Douglas, A. E. (2013). Microbial brokers of insect-plant interactions revisited. *Journal of Chemical Ecology*, *39*(7), 952–961.

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

Engelbrektson, A., Hubbard, C. G., Tom, L. M., Boussina, A., Jin, Y. T., Wong, H., Piceno, Y. M., Carlson, H. K., Conrad, M. E., Anderson, G., & Coates, J. D. (2014). Inhibition of microbial sulfate reduction in a flow-through column system by (per)chlorate treatment. *Frontiers in Microbiology*, *5*, 315.

Engel, P., & Moran, N. A. (2013). The gut microbiota of insects - diversity in structure and function. *FEMS Microbiology Reviews*, *37*(5), 699–735.

Fan, J., Crooks, C., Creissen, G., Hill, L., Fairhurst, S., Doerner, P., & Lamb, C. (2011). Pseudomonas sax genes overcome aliphatic isothiocyanate-mediated non-host resistance in Arabidopsis. *Science*, *331*(6021), 1185–1188.

Freeland, W. J., & Janzen, D. H. (1974). Strategies in Herbivory by Mammals: The Role of Plant Secondary Compounds. *The American Naturalist*, *108*(961), 269–289.

Futuyma, D. J., & Agrawal, A. A. (2009). Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(43), 18054–18061.

Gloss, A. D., Vassão, D. G., Hailey, A. L., Nelson Dittrich, A. C., Schramm, K., Reichelt, M., Rast, T. J., Weichsel, A., Cravens, M. G., Gershenzon, J., Montfort, W. R., & Whiteman, N. K. (2014). Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the drosophilidae. *Molecular Biology and Evolution*, *31*(9), 2441–2456.

Grenier, T., & Leulier, F. (2020). How commensal microbes shape the physiology of Drosophila melanogaster. *Current Opinion in Insect Science*, *41*, 92–99.

Hammer, T. J., & Bowers, M. D. (2015). Gut microbes may facilitate insect herbivory of chemically defended plants. *Oecologia*, *179*(1), 1–14.

Hansen, A. K., & Moran, N. A. (2014). The impact of microbial symbionts on host plant utilization by herbivorous insects. *Molecular Ecology*, *23*(6), 1473–1496.

Heys, C., Fisher, A. M., Dewhurst, A. D., Lewis, Z., & Lizé, A. (2021). Exposure to foreign gut microbiota can facilitate rapid dietary shifts. *Scientific Reports*, *11*(1), 16791.

Huang, Y., Niu, B., Gao, Y., Fu, L., & Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* , *26*(5), 680–682.

Ja, W. W., Carvalho, G. B., Mak, E. M., de la Rosa, N. N., Fang, A. Y., Liong, J. C., Brummel, T., & Benzer, S. (2007). Prandiology of *Drosophila* and the CAFE assay. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8253–8256.

Kang, K., Panzano, V. C., Chang, E. C., Ni, L., Dainis, A. M., Jenkins, A. M., Regna, K., Muskavitch, M. A. T., & Garrity, P. A. (2011). Modulation of TRPA1 thermal sensitivity enables sensory discrimination in Drosophila. *Nature*, *481*(7379), 76–80.

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.

Lichtenstein, E. P., Strong, F. M., & Morgan, D. G. (1962). Naturally Occurring Insecticides, Identification of 2-Phenylethylisothiocyanate As an Insecticide Occurring Naturally in the Edible Part of Turnips. *Journal of Agricultural and Food Chemistry*, *10*(1), 30–33.

Ludington, W. B., & Ja, W. W. (2020). Drosophila as a model for the gut microbiome. *PLoS Pathogens*, *16*(4), e1008398.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Del Rio, T. G., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G., & Dangl, J. L. (2012). Defining the core Arabidopsis thaliana root microbiome. *Nature*, *488*(7409), 86–90.

Mangiafico, S., & Mangiafico, M. S. (2017). Package "rcompanion." *Cran Repos*, *20*, 1–71.

Mason, C. J. (2020). Complex Relationships at the Intersection of Insect Gut Microbiomes and Plant Defenses. *Journal of Chemical Ecology*, *46*(8), 793–807.

Mason, C. J., Couture, J. J., & Raffa, K. F. (2014). Plant-associated bacteria degrade defense chemicals and reduce their adverse effects on an insect defoliator. *Oecologia*, *175*(3), 901–910.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2011). The CIPRES science gateway: a community resource for phylogenetic analyses. *Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery*, Article Article 41.

Mitter, C., Farrell, B., & Wiegmann, B. (1988). The Phylogenetic Study of Adaptive Zones: Has Phytophagy Promoted Insect Diversification? *The American Naturalist*, *132*(1), 107–128.

Murphy, K. C. (1998). Use of bacteriophage lambda recombination functions to promote gene replacement in Escherichia coli. *Journal of Bacteriology*, *180*(8), 2063–2071.

O'Connor, T. K., Humphrey, P. T., Lapoint, R. T., Whiteman, N. K., & O'Grady, P. M. (2014). Microbial interactions and the ecology and evolution of Hawaiian Drosophilidae. *Frontiers in Microbiology*, *5*, 616.

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414.

Peláez, J. N., Gloss, A. D., Goldman-Huertas, B., Kim, B., Lapoint, R. T., Pimentel-Solorio, G., Verster, K. I., Aguilar, J. M., Nelson Dittrich, A. C., Singhal, M., Suzuki, H. C., Matsunaga, T., Armstrong, E. E., Charboneau, J. L. M., Groen, S. C., Hembry, D. H., Ochoa, C. J., O'Connor, T. K., Prost, S., … Whiteman, N. K. (2023). Evolution of chemosensory and detoxification gene families across herbivorous Drosophilidae. *G3* . https://doi.org/10.1093/g3journal/jkad133

Peláez, J. N., Gloss, A. D., Ray, J. F., Chaturvedi, S., Haji, D., Charboneau, J. L. M., Verster, K. I., & Whiteman, N. K. (2022). Evolution and genomic basis of the plant-

penetrating ovipositor: a key morphological trait in herbivorous Drosophilidae. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1986), 20221938.

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schoonhoven, L., van Loon, J. J. A., & Dicke, M. (2005). *Insect-Plant Biology*. Oxford University Press.

Scott, K. (2018). Gustatory Processing in Drosophila melanogaster. *Annual Review of Entomology*, *63*, 15–30.

Shukla, S. P., & Beran, F. (2020). Gut microbiota degrades toxic isothiocyanates in a flea beetle pest. *Molecular Ecology*, *29*(23), 4692–4705.

Southwood, T. R. E. (1972). The insect/plant relationship--an evolutionary perspective. *Royal Entomological Society London Symposium*, *6*, 3–30.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* , *30*(9), 1312–1313.

Tang, C. S., Bhothipaksa, K., & Frank, H. A. (1972). Bacterial degradation of benzyl isothiocyanate. *Applied Microbiology*, *23*(6), 1145–1148.

Textor, S., & Gershenzon, J. (2009). Herbivore induction of the glucosinolate–myrosinase defense system: major trends, biochemical bases and ecological significance. *Phytochemistry Reviews*, *8*(1), 149–170.

van den Bosch, T. J. M., Niemi, O., & Welte, C. U. (2020). Single gene enables plant pathogenic Pectobacterium to overcome host-specific chemical defence. *Molecular Plant Pathology*, *21*(3), 349–359.

van den Bosch, T. J. M., Tan, K., Joachimiak, A., & Welte, C. U. (2018). Functional Profiling and Crystal Structures of Isothiocyanate Hydrolases Found in Gut-Associated and Plant-Pathogenic Bacteria. *Applied and Environmental Microbiology*, *84*(14). https://doi.org/10.1128/AEM.00478-18

van den Bosch, T. J. M., & Welte, C. U. (2017). Detoxifying symbionts in agriculturally important pest insects. *Microbial Biotechnology*, *10*(3), 531–540.

Wei, J., Segraves, K. A., Li, W.-Z., Yang, X.-K., & Xue, H.-J. (2020). Gut bacterial communities and their contribution to performance of specialist Altica flea beetles. *Microbial Ecology*, *80*(4), 946–959.

Welte, C. U., de Graaf, R. M., van den Bosch, T. J. M., Op den Camp, H. J. M., van Dam, N. M., & Jetten, M. S. M. (2016). Plasmids from the gut microbiome of cabbage root fly larvae encode SaxA that catalyses the conversion of the plant toxin 2-phenylethyl isothiocyanate. *Environmental Microbiology*, *18*(5), 1379–1390.

Welte, C. U., Rosengarten, J. F., de Graaf, R. M., & Jetten, M. S. M. (2016). SaxA-Mediated Isothiocyanate Metabolism in Phytopathogenic Pectobacteria. *Applied and Environmental Microbiology*, *82*(8), 2372–2379.

Whiteman, N. K., Gloss, A. D., Sackton, T. B., Groen, S. C., Humphrey, P. T., Lapoint, R. T., Sønderby, I. E., Halkier, B. A., Kocks, C., Ausubel, F. M., & Pierce, N. E. (2012). Genes involved in the evolution of herbivory by a leaf-mining, Drosophilid fly. *Genome Biology and Evolution*, *4*(9), 900–916.

Whiteman, N. K., Groen, S. C., Chevasco, D., Bear, A., Beckwith, N., Gregory, T. R., Denoux, C., Mammarella, N., Ausubel, F. M., & Pierce, N. E. (2011). Mining the plant-herbivore interface with a leafmining Drosophila of Arabidopsis. *Molecular Ecology*, *20*(5), 995–1014.

Wiens, J. J., Lapoint, R. T., & Whiteman, N. K. (2015). Herbivory increases diversification across insect clades. *Nature Communications*, *6*, 8370.

References Chapter 3

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.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* , *34*(17), i884–i890.

Chung, H., Bogwitz, M. R., McCart, C., Andrianopoulos, A., Ffrench-Constant, R. H., Batterham, P., & Daborn, P. J. (2007). Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the Drosophila melanogaster insecticide resistance gene Cyp6g1. *Genetics*, *175*(3), 1071–1077.

Daborn, P., Boundy, S., Yen, J., Pittendrigh, B., & ffrench-Constant, R. (2001). DDT resistance in Drosophila correlates with Cyp6g1 over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Molecular Genetics and Genomics: MGG*, *266*(4), 556–563.

Daborn, P. J., Yen, J. L., Bogwitz, M. R., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P., Feyereisen, R., Wilson, T. G., & ffrench-Constant, R. H. (2002). A single p450 allele associated with insecticide resistance in Drosophila. *Science*, *297*(5590), 2253–2256.

De Bie, T., Cristianini, N., Demuth, J. P., & Hahn, M. W. (2006). CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* , *22*(10), 1269–1271.

Dixit, S., Widemann, E., Bensoussan, N., Salehipourshirazi, G., Bruinsma, K., Milojevic, M., Shukla, A., Romero, L. C., Zhurov, V., Bernards, M. A., Chruszcz, M., Grbić, M., & Grbić, V. (2022). β-Cyanoalanine synthase protects mites against Arabidopsis defenses. *Plant Physiology*, *189*(4), 1961–1975.

Dobzhansky, T. (1937). *Genetics and the origin of species* (Vol. 323). Columbia University Press.

Drosophila 12 Genomes Consortium, Clark, A. G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., Kaufman, T. C., Kellis, M., Gelbart, W., Iyer, V. N., Pollard, D. A., Sackton, T. B., Larracuente, A. M., Singh, N. D., Abad, J. P., Abt, D. N., Adryan, B., Aguade, M., … MacCallum, I. (2007). Evolution of genes and genomes on the Drosophila phylogeny. *Nature*, *450*(7167), 203–218.

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

Frank, M. R., & Fogleman, J. C. (1992). Involvement of cytochrome P450 in host-plant utilization by Sonoran Desert Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(24), 11998–12002.

Futuyma, D. J., & Agrawal, A. A. (2009). Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(43), 18054–18061.

Gloss, A. D., Vassão, D. G., Hailey, A. L., Nelson Dittrich, A. C., Schramm, K., Reichelt, M., Rast, T. J., Weichsel, A., Cravens, M. G., Gershenzon, J., Montfort, W. R., & Whiteman, N. K. (2014). Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the drosophilidae. *Molecular Biology and Evolution*, *31*(9), 2441–2456.

Gordon, H. T. (1961). Nutritional factors in insect resistance to chemicals. *Annual Review of Entomology*, *6*, 27–54.

Hardy, N. B., Peterson, D. A., Ross, L., & Rosenheim, J. A. (2018). Does a plant-eating insect's diet govern the evolution of insecticide resistance? Comparative tests of the pre-adaptation hypothesis. *Evolutionary Applications*, *11*(5), 739–747.

Harrop, T. W. R., Sztal, T., Lumb, C., Good, R. T., Daborn, P. J., Batterham, P., & Chung, H. (2014). Evolutionary changes in gene expression, coding sequence and copy-number at the Cyp6g1 locus contribute to resistance to multiple insecticides in Drosophila. *PloS One*, *9*(1), e84879.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522.

Hodges, R. E., & Minich, D. M. (2015). Modulation of Metabolic Detoxification Pathways Using Foods and Food-Derived Components: A Scientific Review with Clinical Application. *Journal of Nutrition and Metabolism*, *2015*, 760689.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research*, *36*(Web Server issue), W5–W9.

Joussen, N., Heckel, D. G., Haas, M., Schuphan, I., & Schmidt, B. (2008). Metabolism of imidacloprid and DDT by P450 CYP6G1 expressed in cell cultures of Nicotiana tabacum suggests detoxification of these insecticides in Cyp6g1-overexpressing strains of Drosophila melanogaster, leading to resistance. *Pest Management Science*, *64*(1), 65–73.

Kim, B. Y., Wang, J. R., Miller, D. E., Barmina, O., Delaney, E., Thompson, A., Comeault, A. A., Peede, D., D'Agostino, E. R., Pelaez, J., Aguilar, J. M., Haji, D., Matsunaga, T., Armstrong, E. E., Zych, M., Ogawa, Y., Stamenković-Radak, M., Jelić, M., Veselinović, M. S., … Petrov, D. A. (2021). Highly contiguous assemblies of 101 drosophilid genomes. *eLife*, *10*. https://doi.org/10.7554/eLife.66405

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915.

Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2019). HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, *37*(1), 295–299.

Labandeira, C. C., & Currano, E. D. (2013). The Fossil Record of Plant-Insect Dynamics. *Annual Review of Earth and Planetary Sciences*, *41*(Volume 41, 2013), 287–311.

Lapoint, R. T., O'Grady, P. M., & Whiteman, N. K. (2013). Diversification and dispersal of the Hawaiian Drosophilidae: the evolution of Scaptomyza. *Molecular Phylogenetics and Evolution*, *69*(1), 95–108.

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296.

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

Li, X., Baudry, J., Berenbaum, M. R., & Schuler, M. A. (2004). Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 2939–2944.

Li, X., Berenbaum, M. R., & Schuler, M. A. (2002). Plant allelochemicals differentially regulate Helicoverpa zea cytochrome P450 genes. *Insect Molecular Biology*, *11*(4), 343–351.

Li, X., Schuler, M. A., & Berenbaum, M. R. (2007). Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annual Review of Entomology*, *52*, 231–253.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550.

Löytynoja, A. (2014). Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology* , *1079*, 155–170.

Mackay, T. F. (1989). Transposable elements and fitness in Drosophila melanogaster. *Genome / National Research Council Canada = Genome / Conseil National de Recherches Canada*, *31*(1), 284–295.

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). https://doi.org/10.1093/molbev/msab362

Mérel, V., Boulesteix, M., Fablet, M., & Vieira, C. (2020). Transposable elements in Drosophila. *Mobile DNA*, *11*, 23.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534.

Naser-Khdour, S., Minh, B. Q., & Lanfear, R. (2021). Assessing Confidence in Root Placement on Phylogenies: An Empirical Study Using Non-Reversible Models for Mammals. In *bioRxiv* (p. 2020.07.31.230144). https://doi.org/10.1101/2020.07.31.230144

Nelson, D. R. (2018). Cytochrome P450 diversity in the tree of life. *Biochimica et Biophysica Acta: Proteins and Proteomics*, *1866*(1), 141–154.

Peláez, J. N., Gloss, A. D., Goldman-Huertas, B., Kim, B., Lapoint, R. T., Pimentel-Solorio, G., Verster, K. I., Aguilar, J. M., Nelson Dittrich, A. C., Singhal, M., Suzuki, H. C., Matsunaga, T., Armstrong, E. E., Charboneau, J. L. M., Groen, S. C., Hembry, D. H., Ochoa, C. J., O'Connor, T. K., Prost, S., … Whiteman, N. K. (2023). Evolution of chemosensory and detoxification gene families across herbivorous Drosophilidae. *G3* . https://doi.org/10.1093/g3journal/jkad133

Peláez, J. N., Gloss, A. D., Ray, J. F., Chaturvedi, S., Haji, D., Charboneau, J. L. M., Verster, K. I., & Whiteman, N. K. (2022). Evolution and genomic basis of the plant-penetrating ovipositor: a key morphological trait in herbivorous Drosophilidae. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1986), 20221938.

Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E., & Zanini, F. (2022). Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* , *38*(10), 2943–2945.

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Salehipourshirazi, G., Bruinsma, K., Ratlamwala, H., Dixit, S., Arbona, V., Widemann, E., Milojevic, M., Jin, P., Bensoussan, N., Gómez-Cadenas, A., Zhurov, V., Grbic, M., & Grbic, V. (2021). Rapid specialization of counter defenses enables two-spotted spider mite to adapt to novel plant hosts. *Plant Physiology*, *187*(4), 2608–2622.

Sánchez-Bayo, F. (2021). Indirect Effect of Pesticides on Insects and Other Arthropods. *Toxics*, *9*(8). https://doi.org/10.3390/toxics9080177

Schlenke, T. A., & Begun, D. J. (2004). Strong selective sweep associated with a transposon insertion in Drosophila simulans. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(6), 1626–1631.

Schmidt, J. M., Good, R. T., Appleton, B., Sherrard, J., Raymant, G. C., Bogwitz, M. R., Martin, J., Daborn, P. J., Goddard, M. E., Batterham, P., & Robin, C. (2010). Copy number variation and transposable elements feature in recent, ongoing adaptation at the Cyp6g1 locus. *PLoS Genetics*, *6*(6), e1000998.

Schuler, M. A. (2011). P450s in plant-insect interactions. *Biochimica et Biophysica Acta*, *1814*(1), 36–45.

Scott, J. G. (1999). Cytochromes P450 and insecticide resistance. *Insect Biochemistry and Molecular Biology*, *29*(9), 757–777.

Smith, M. D., Wertheim, J. O., Weaver, S., Murrell, B., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution*, *32*(5), 1342–1353.

Southwood, T. R. E. (1972). The insect/plant relationship--an evolutionary perspective. *Royal Entomological Society London Symposium*, *6*, 3–30.

Sparks, T. C., DeBoer, G. J., Wang, N. X., Hasler, J. M., Loso, M. R., & Watson, G. B. (2012). Differential metabolism of sulfoximine and neonicotinoid insecticides by Drosophila melanogaster monooxygenase CYP6G1. *Pesticide Biochemistry and Physiology*, *103*(3), 159–165.

Stevens, J. L., Snyder, M. J., Koener, J. F., & Feyereisen, R. (2000). Inducible P450s of the CYP9 family from larval Manduca sexta midgut. *Insect Biochemistry and Molecular Biology*, *30*(7), 559–568.

Umetsu, N., & Shirai, Y. (2020). Development of novel pesticides in the 21st century. *Journal of Pesticide Science*, *45*(2), 54–74.

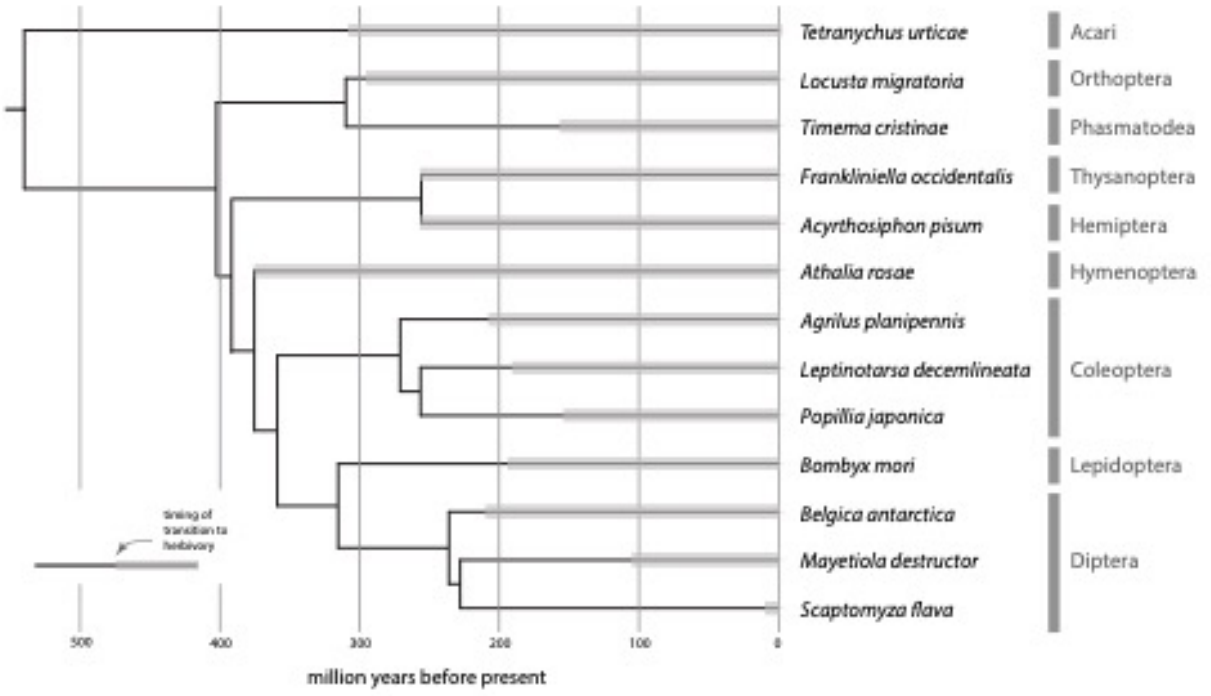
Vázquez-Gómez, G., Sánchez-Santos, A., Vázquez-Medrano, J., Quintanar-Zúñiga, R., Monsalvo-Reyes, A. C., Piedra-Ibarra, E., Dueñas-García, I. E., Castañeda-Partida, L., Graf, U., & Heres-Pulido, M. E. (2010). Sulforaphane modulates the expression of Cyp6a2 and Cyp6g1 in larvae of the ST and HB crosses of the Drosophila wing spot test and is genotoxic in the ST cross. *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, *48*(12), 3333–3339.

Whiteman, N. K., Gloss, A. D., Sackton, T. B., Groen, S. C., Humphrey, P. T., Lapoint, R. T., Sønderby, I. E., Halkier, B. A., Kocks, C., Ausubel, F. M., & Pierce, N. E. (2012). Genes involved in the evolution of herbivory by a leaf-mining, Drosophilid fly. *Genome Biology and Evolution*, *4*(9), 900–916.

Appendix 1

Timing of insect transitions to herbivory

To estimate the timing of transitions to herbivory in lineages with members that have had their genomes sequenced, we searched for exemplar clades of herbivorous species with reference genomes on InsectBase 2.0 in July of 2022 (Mei et al. 2022). From information available in the literature, we categorized them as either obligate herbivore (having at least one life stage where it feeds exclusively on living plant tissues) or facultative herbivore (feeding frequently on plants, but not having obligate herbivory stages). We focused on leaf-feeding insects, and excluded root-feeders, frugivores and granivores. We used stem ages and divergence times (Supplementary Table 1) of herbivorous lineages with at least one representative genome assembly from the literature (Ahrens et al., 2014; McKenna et al., 2019; Misof et al., 2014; Wiegmann et al., 2011; Wiens et al., 2015). We inferred that the most recent common ancestor of an herbivorous clade was also herbivorous under maximum parsimony and therefore considered these stem ages as estimated dates for origins of herbivory. We visualized the phylogenetic relationships among reference herbivores with iTOL v6 (Letunic & Bork, 2021), and included the herbivorous spider mite (*Tetranychus urticae*) as an outgroup (Rota-Stabelli, Daley, and Pisani 2013; Xue et al. 2017).



Supplementary Figure 1: The timing of transitions to herbivory, inferred from the reported diet niche of related families and orders, in the ancestors of exemplar clades of terrestrial herbivorous arthropods with sequenced genomes.

Node	Estimated divergence time (MYA)	Data source	Divergence estimation method
Scaptomyza - Mayetiola	228	Figure S3, Weigmann et al. (2011) [rough estimation]	Maximum Likelihood
Belgica - Scaptomyza/Mayetiola	236	Figure S3, Weigmann et al. (2011) [rough estimation]	Maximum Likelihood
Lepidoptera - Diptera	315.3	Figure S2, Wiens et al. (2015)	Bayesian
Popillia - Leptinotarsa	256.3	Figure S13, McKenna et al. (2019)	Maximum Likelihood
Agrilus - Popillia/Leptinotarsa	270.9	Figure S13, McKenna et al. (2019)	Maximum Likelihood
Hymenoptera - Col/Lep/Dip	375.4	Figure S2, Wiens et al. (2015)	Bayesian
Thysanoptera - Hemiptera	255.8	Figure S2, Wiens et al. (2015)	Bayesian
Thy/Hem - Hym/Col/Lep/Dip	391.6	Figure S2, Wiens et al. (2015)	Bayesian
Orthoptera - Phasmatodae	309.8	Figure S2, Wiens et al. (2015)	Bayesian
Ort/Pha - Other insects	402.4	Figure S2, Wiens et al. (2015)	Bayesian
Tetranychus - Insects	540	Figure 1, Rota-Stabelli et al. (2013) [rough estimation]	Bayesian

Supplementary Table 1: References and estimation methods for divergence times presented in Supplementary Figure 1.

Yeast media ingredients

UC Berkeley yeast media ingredients:

water, agar, cornmeal, yeast, sucrose, sodium tartrate (4 hydrate), $CaCl_2$ (2 hydrate), molasses, tegosept in 95% EtOH, and propionic acid

Bloomington yeast media ingredients:

water, cornmeal, light corn syrup, malt extract, yeast, soy flour, agar, antimicrobial agents

Sterol extraction and processing

We extracted samples using equal volumes (8 milliliters) of chloroform, 70% methanol, and water. Following a 12 hour incubation, the chloroform layer containing the free and base hydrolysable sterols was removed and evaporated to a reduced volume (less than 100 μl). We set aside the 70% methanol:water layer containing glycosylated sterols for additional processing and sterol identification.

We resuspended the reduced volume chloroform layer (less than 100 μl but not dryness) in 5 mls of hexane and 5 mls of 70% methanol/water combined. We removed the original hexane layer following separation of the phases. We added a second 5 ml volume of hexane to the original 70% methanol/water layer and mixed vigorously. After separation, we removed the second hexane layer, combined the two hexane volumes, and evaporated to a reduced volume – less than 100 ul.

To release base hydrolysable sterol esters, we resuspended the hexane fraction in 8 ml of 70% methanol/water containing 5% KOH w/v (with sonication) and incubated at 70°C in a water bath for 1.5 hours with constant shaking (Feng et al. 2015). After base saponification, we added 3 ml of water and 5 ml of water equilibrated hexane, shook vigorously and vortexed for 1 minute. We then left samples for at least 6 hours, until the layers separated, or centrifuged for 2 minutes to assist in phase separation (MSE GT-2 centrifuge equipped with a swinging bucket rotor – at 1,500 rpms - VWR). We removed the hexane layer (containing the released sterols). We added 4 ml of hexane to the remaining methanol/water, repeated the above steps, and again removed the hexane layer. Using the above separation protocols, we combined the two hexane volumes and backwashed to neutrality against an equal volume of water, with neutrality determined by PHydrion Mikro Ion paper (Micro Essential Laboratories, Brooklyn, NY). We concentrated the neutral hexane fraction by evaporating to a minimal volume (less than 100 μl). All of the described procedures used solvents that were pre-equilibrated to

the paired solvent. We followed the procedure described in Feng et al., 2015 to release the glycosylated sterols.

	A) Full statistical results				
	Gut/body	Cephalo-skeleton/ body	Mouth hook height/length	Mouth hook tooth #	Mouth hook toothsize
Df	3	3	3	3	3
Df, residuals	36	31	28	33	33
Sum Sq	6.946	0.006	0.584	0.003	217.80
Sum Sq, residuals	3.358	0.002	0.299	1.85e-4	27.12
Mean Sq	2.315	0.002	0.195	1.0e-03	72.6
Mean Sq, residuals	0.0933	5.46e-5	0.011	5.6e-06	0.82
F-value	24.82	34.7	18.23	178	88.33
Pr(>F)	<0.001	<0.001	<0.001	<0.001	<0.001

	B) Tukey's HSD species comparisons				
D. melanogaster / *S. hsui*	<0.001	0.580	0.001	<0.001	<0.001
D. melanogaster / *S. pallida*	<0.001	<0.001	0.544	<0.001	<0.001
D. melanogaster / *S.flava*	<0.001	<0.001	<0.001	<0.001	<0.001
S. pallida / *S. hsui*	0.955	<0.001	0.033	0.019	0.454
S. hsui / *S. flava*	0.051	<0.001	0.086	0.028	<0.001
S. pallida / *S. flava*	0.134	0.367	<0.001	<0.001	<0.001

Supplementary table 2. A) full statistical output of ANOVA comparing measurements of body parts between *D. melanogaster, S. hsui, S. pallida, S. flava* B) P values of morphological measurements adjusted for multiple comparisons with Tukey's honestly significant difference test

Supplementary Citations

Ahrens, Dirk, Julia Schwarzer, and Alfried P. Vogler. 2014. “The Evolution of Scarab Beetles Tracks the Sequential Rise of Angiosperms and Mammals.” *Proceedings. Biological Sciences / The Royal Society* 281 (1791): 20141470.

Feng, Simin, Songbai Liu, Zisheng Luo, and Kaichen Tang. 2015. “Direct Saponification Preparation and Analysis of Free and Conjugated Phytosterols in Sugarcane (Saccharum Officinarum L.) by Reversed-Phase High-Performance Liquid Chromatography.” *Food Chemistry* 181 (August): 9–14.

Letunic, Ivica, and Peer Bork. 2021. “Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation.” *Nucleic Acids Research* 49 (W1): W293–96.

McKenna, Duane D., Seunggwan Shin, Dirk Ahrens, Michael Balke, Cristian Beza-Beza, Dave J. Clarke, Alexander Donath, et al. 2019. “The Evolution and Genomic Basis of Beetle Diversity.” *Proceedings of the National Academy of Sciences of the United States of America* 116 (49): 24729–37.

Mei, Yang, Dong Jing, Shenyang Tang, Xi Chen, Hao Chen, Haonan Duanmu, Yuyang Cong, et al. 2022. “InsectBase 2.0: A Comprehensive Gene Resource for Insects.” *Nucleic Acids Research* 50 (D1): D1040–45.

Misof, Bernhard, Shanlin Liu, Karen Meusemann, Ralph S. Peters, Alexander Donath, Christoph Mayer, Paul B. Frandsen, et al. 2014. “Phylogenomics Resolves the Timing and Pattern of Insect Evolution.” *Science* 346 (6210): 763–67.

Rota-Stabelli, Omar, Allison C. Daley, and Davide Pisani. 2013. “Molecular Timetrees Reveal a Cambrian Colonization of Land and a New Scenario for Ecdysozoan Evolution.” *Current Biology: CB* 23 (5): 392–98.

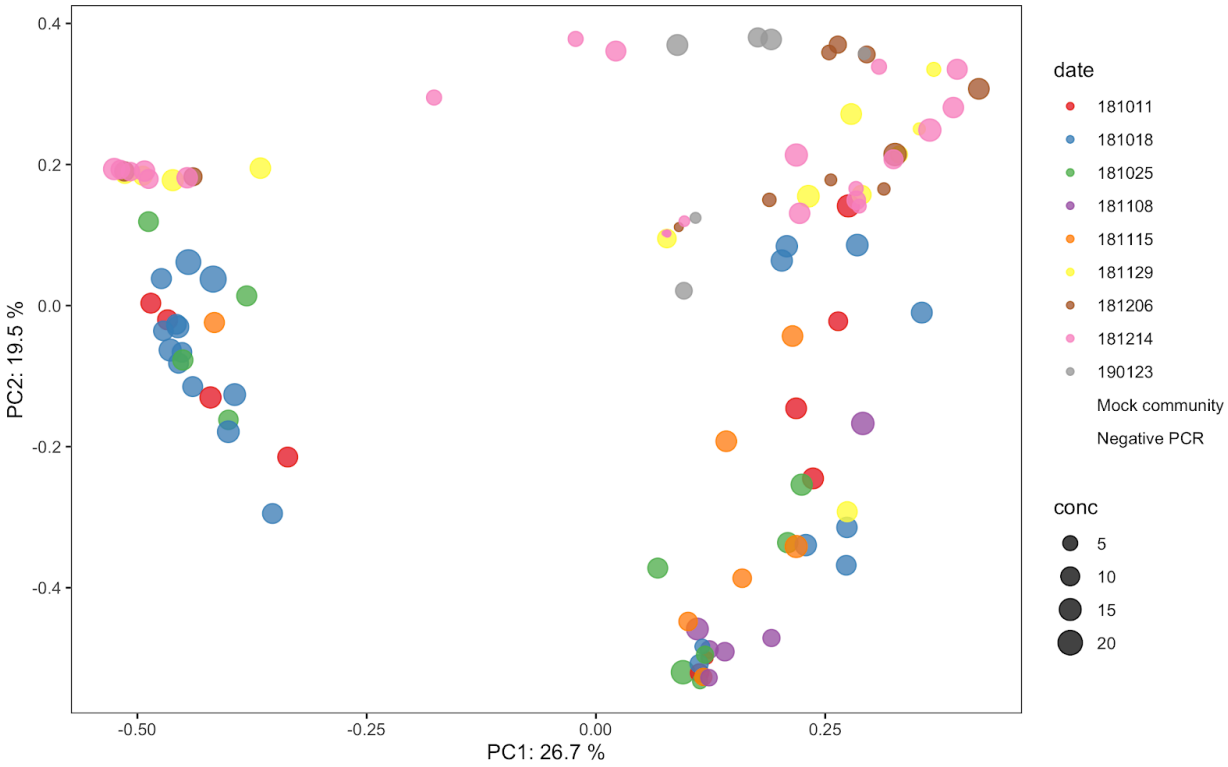
Wiegmann, Brian M., Michelle D. Trautwein, Isaac S. Winkler, Norman B. Barr, Jung-Wook Kim, Christine Lambkin, Matthew A. Bertone, et al. 2011. “Episodic Radiations in the Fly Tree of Life.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (14): 5690–95.

Wiens, John J., Richard T. Lapoint, and Noah K. Whiteman. 2015. “Herbivory Increases Diversification across Insect Clades.” *Nature Communications* 6 (September): 8370.

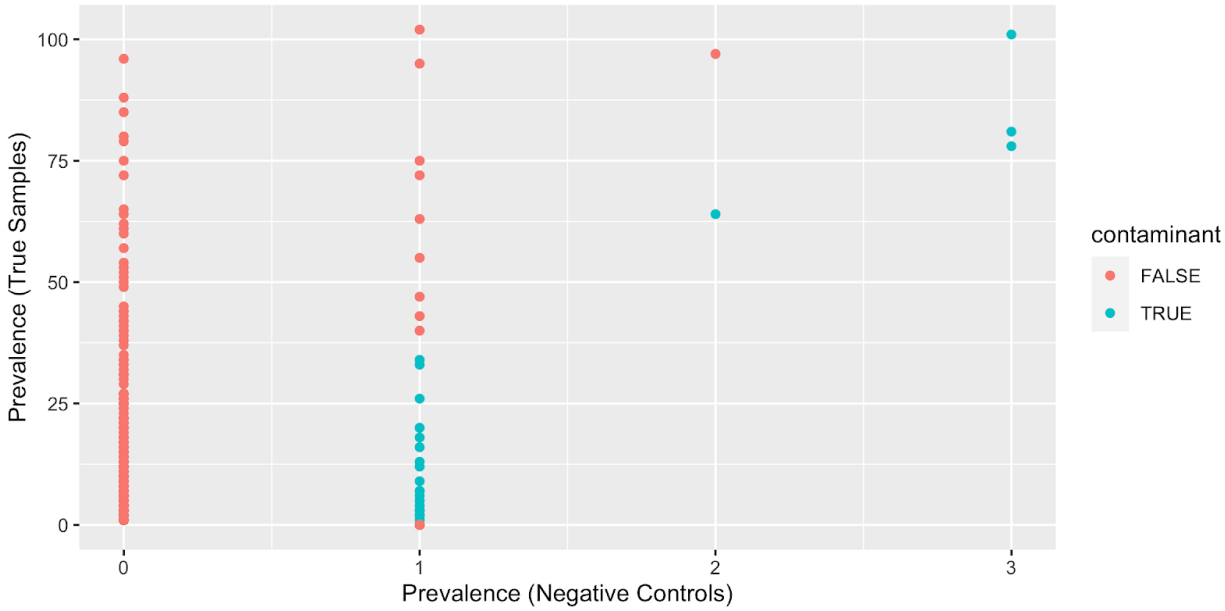
Xue, Xiao-Feng, Yan Dong, Wei Deng, Xiao-Yue Hong, and Renfu Shao. 2017. “The Phylogenetic Position of Eriophyoid Mites (Superfamily Eriophyoidea) in Acariformes Inferred from the Sequences of Mitochondrial Genomes and Nuclear Small Subunit (18S) RRNA Gene.” *Molecular Phylogenetics and Evolution* 109: 271–82.

Appendix 2

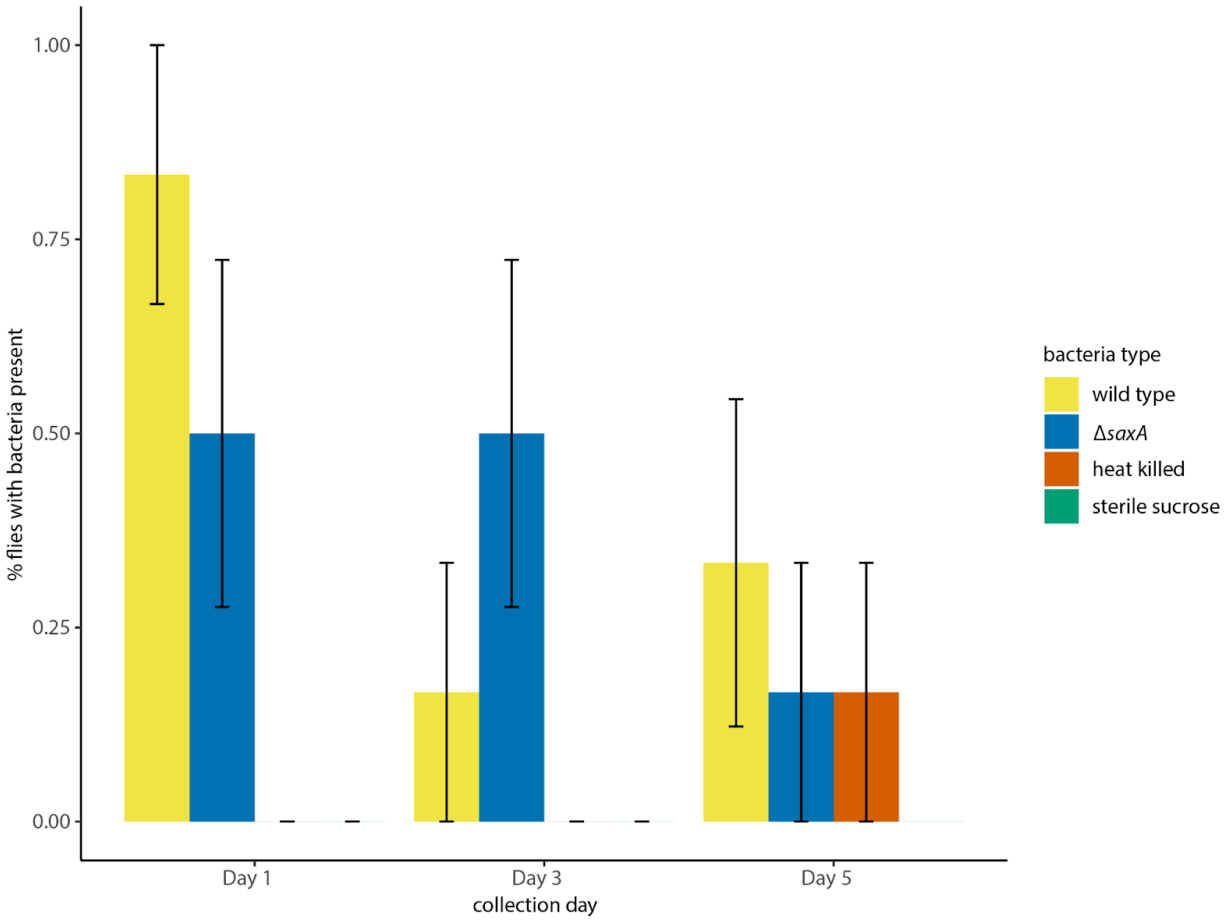
Supplementary figures



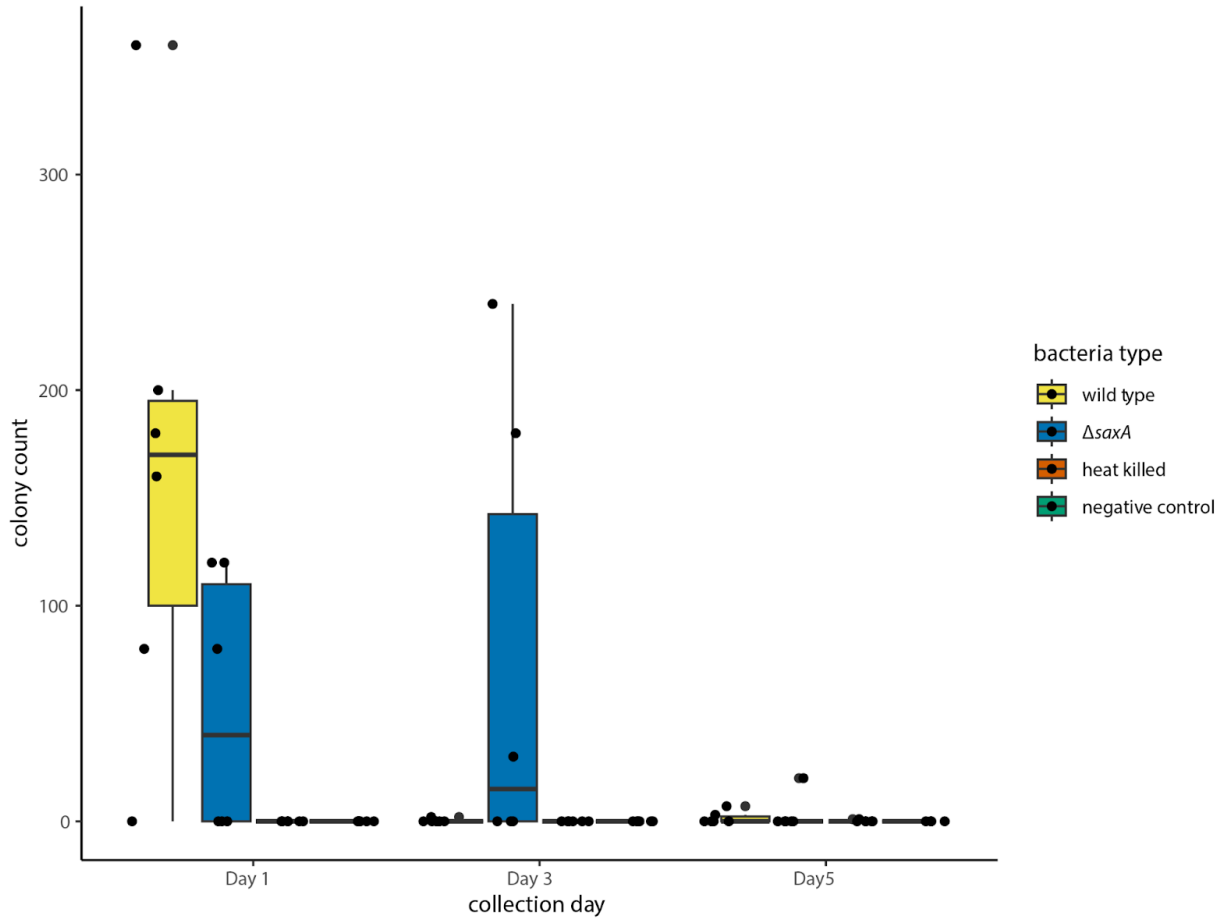
Supplementary Figure 1 : PCoA of Bray-Curtis distances using DADA2-denoised read counts across all samples. No additional filtering was performed. “Date” refers to the date the sample was collected and the concentration is what was reported in the metadata.



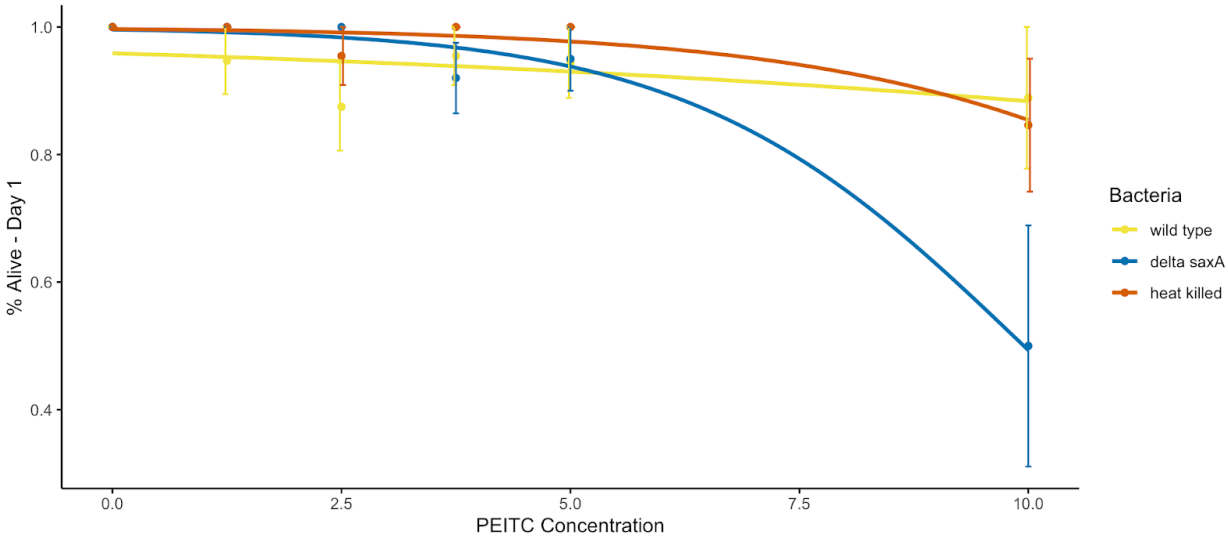
Supplementary Figure 2: Relationship between the prevalence of individual ASVs in controls compared to non-controls. Contaminant ASVs were identified using Decontam's prevalence-based contaminant filtering method with a threshold of 0.5. See "List of contaminant taxa identified with Decontam (threshold 0.5)" at the end of this document.



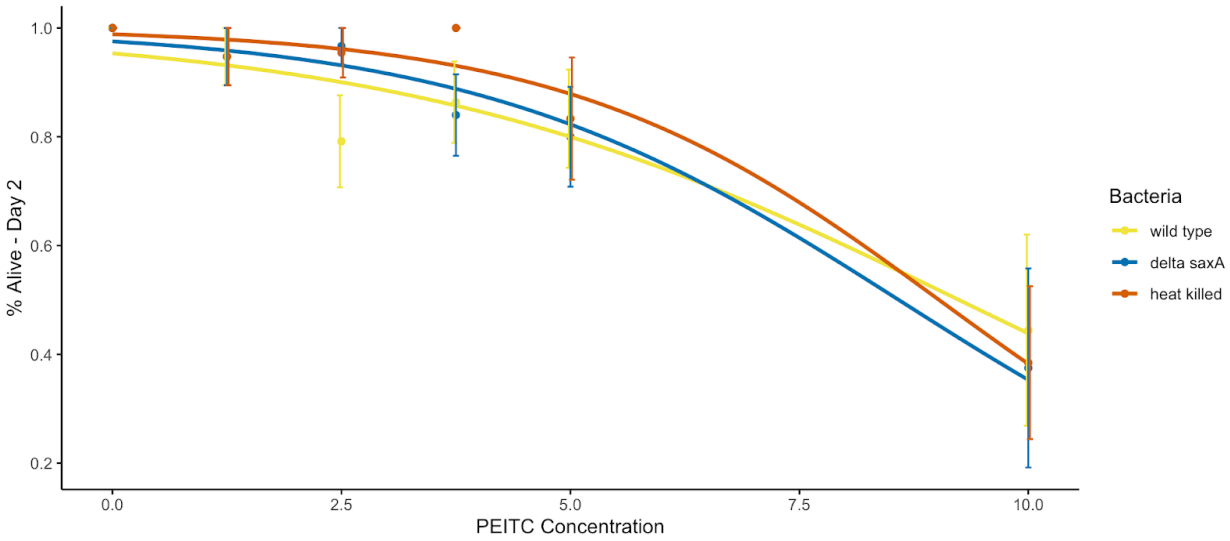
Supplementary Figure 3: Introduced bacteria persistence in the gut of *S. flava* across a five day interval.



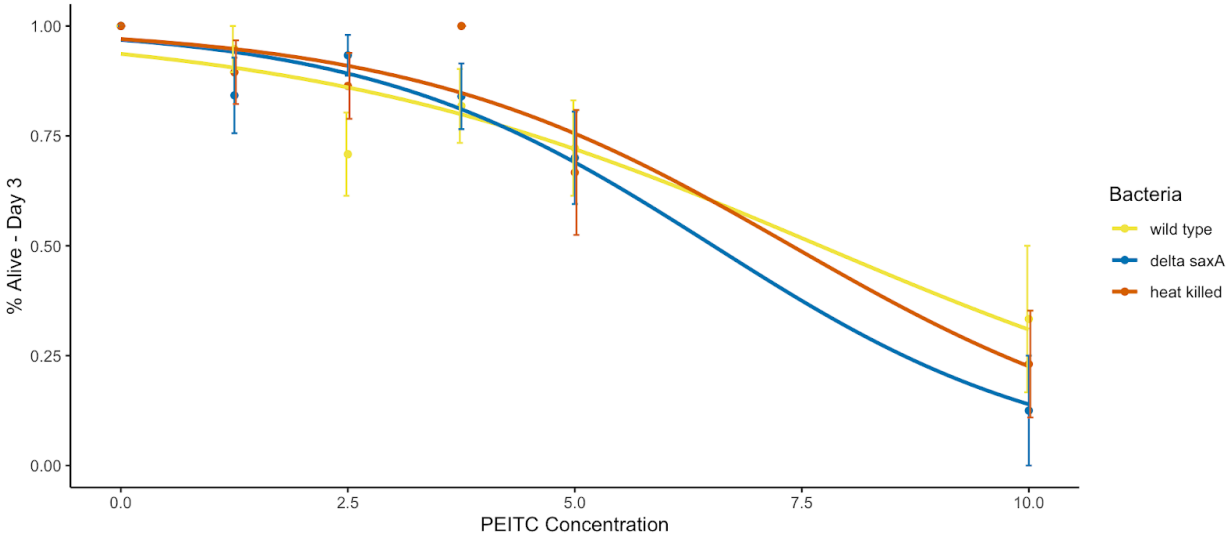
Supplementary Figure 4 - Amount of introduced bacteria that persists over time is variable Supplementary



Supplementary Figure 5 - CAFE survival curves for day 1.



Supplementary Figure 6 - CAFE survival curves for day 2.



Supplementary Figure 7- CAFE survival curves for day 3.

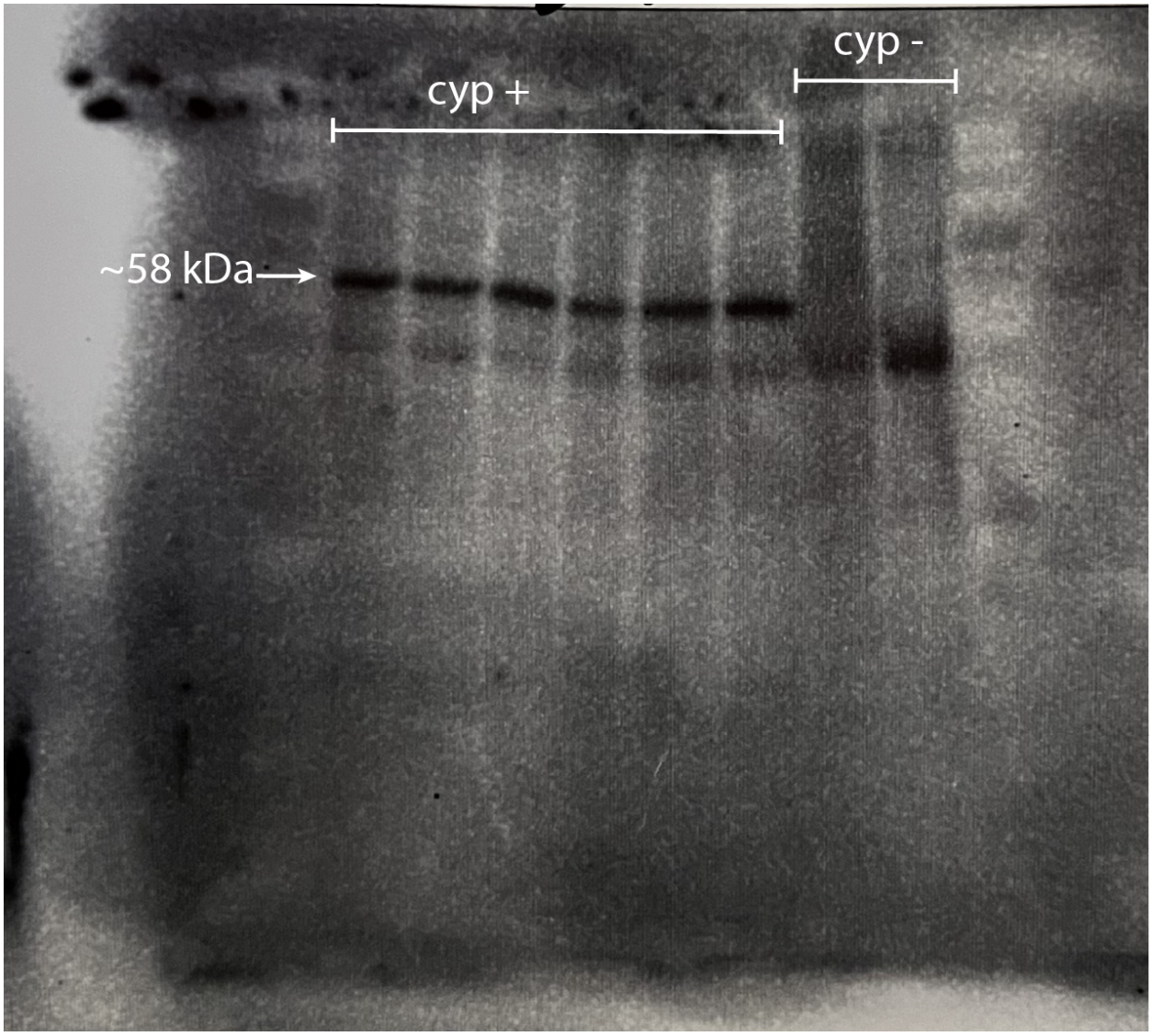
Supplementary Table1 - PCR primers, peptide nucleic acids, and thermocycler conditions

Gene	Primers	Annealing Temperature	Reference
PCR			
col	2183: 5'-CAACATTTATTTTGATTTTTTGG-3' 3041: 5'-TYCATTGCACTAATCTGCCATATTAG-3'		
16S	27f: 5'-AGAGTTTGATCCTGGCTCAG-3' 1492r: 5'-TACCTTGTTACGACTT-3'		
MiSeq			
16S V3-V4	515F: 5'-GTGCCAGCMGCCGCGGTAA-3' 806R: 5'-GGACTACHVHHHTWTCTAAT-3'	50°C	(Parada et al., 2016) (Apprill et al., 2016) (Lundberg et al., 2012)
mPNA	5'-GGCAAGTGTTCTTCGGA-3'		
pPNA	5'-GGCTCAACCCTGGACAG-3'		

Supplementary Table 2 - Amount Eaten -Tukey's HSD statistics				
Comparison	diff	lwr	lwr.1	p.adj
3.75-10	16.49	-42.22	-42.22	0.97
5-10	31.98	-26.04	-26.04	0.61
2.5-10	32.20	-29.93	-29.93	0.67
1.25-10	58.65	-2.04	-2.04	0.06
0-10	106.61	45.92	45.92	1.17E-05
5-3.75	15.49	-29.09	-29.09	0.92
2.5-3.75	15.71	-34.11	-34.11	0.95
1.25-3.75	42.16	-5.84	-5.84	0.12
0-3.75	90.12	42.11	42.11	2.11E-06
2.5-5	0.23	-48.77	-48.77	1
1.25-5	26.67	-20.47	-20.47	0.58
0-5	74.63	27.48	27.48	0.00012
1.25-2.5	26.45	-25.69	-25.69	0.69
0-2.5	74.40	22.27	22.27	0.00076
0-1.25	47.96	-2.44	-2.44	0.073

Supplementary Table 3 - PEITC dose - Tukey's HSD statistics table				
Comparison	diff	lwr	lwr.1	p.adj
1.25-0	93.80	61.06	61.06	1.02E-12
2.5-0	154.53	123.91	123.91	8.77E-13
3.75-0	188.91	157.74	157.74	8.77E-13
5-0	199.71	165.85	165.85	8.77E-13
10 -0	260.32	220.90	220.90	8.77E-13
2.5-1.25	60.73	30.11	30.11	4.34E-07
3.75-1.25	95.12	63.94	63.94	9.33E-13
5-1.25	105.91	72.05	72.05	9.32E-13
10-1.25	166.52	127.10	127.10	8.79E-13
3.75-2.5	34.38	5.43	5.43	0.0096
5-2.5	45.18	13.35	13.35	0.00083
10-2.5	105.78	68.10	68.10	1.18E-12
5-3.75	10.80	-21.56	-21.56	0.93
10-3.75	71.40	33.27	33.27	2.27E-06
5-10	60.61	20.25	20.25	0.00032

Appendix 3



Supplementary Figure 1: Western blot shows cyp6g1 proteins are being produced by S2 cells transfected with purified plasmid containing *cyp6g1* under an actin promoter.

FamilyDesc	Family	MEL	VIR	MOJ	GRI	PAL	HSU	GRA	FLA	ERE	ANA	PSE
Cyp61	Cyp6g1	1	3	1	2	4	2	4	7	1	1	1

Supplementary table 1 : input data for CAFE analysis –*cyp6g1* copy number for each species included in the analysis. Accompanying Newick tree is as follows:

(((ANA:27.80043,(ERE:10.068452,MEL:10.068452):17.731978):14.845397,PSE:42.645827):17.49715,((GRI:33.451265,((((MON:3.77867672,FLA:3.77867672):6.09668628,GRA:9.875363):13.073742,PAL:22.949105):2.112941,HSU:25.062046):8.389219):8.363178,(MOJ:33.962021,VIR:33.962021):7.852422):18.328534)

Supplementary Table 1: Tests for selection

BRANCH	RATES	MAX. DN/DS	UNCORRECTED P-VALUE	CORRECTED P-VALUE
NODE2	1	0.07 (100.00%)	1.00	1
NODE3	1	0.12 (100.00%)	1.00	1
NODE6	2	2.22 (3.75%)	0.40	1
NODE8	1	0.50 (100.00%)	1.00	1
NODE11	2	0.94 (16.38%)	1.00	1
NODE14	2	30.13 (0.79%)	0.13	1
NODE15	2	4.84 (3.14%)	0.18	1
NODE16	2	2.46 (4.45%)	0.20	1
NODE17	1	0.19 (100.00%)	1.00	1
NODE18	2	>1000 (1.19%)	9.23e-6	0.0005
NODE19	1	0.18 (100.00%)	1.00	1
NODE20	1	0.21 (100.00%)	1.00	1
NODE21	1	1.62 (100.00%)	0.35	1
NODE24	2	431.08 (2.00%)	0.01	0.316
NODE27	1	0.68 (100.00%)	1.00	1
NODE28	2	505.96 (0.87%)	0.03	1
NODE29	2	>1000 (0.24%)	0.02	0.958
NODE32	1	0.26 (100.00%)	1.00	1
NODE36	1	0.22 (100.00%)	1.00	1
NODE37	1	0.62 (100.00%)	1.00	1
NODE41	2	14.89 (0.39%)	0.08	1
NODE45	2	1.79 (3.54%)	0.41	1
NODE46	1	>1000 (100.00%)	0.34	1
NODE48	2	0.73 (11.03%)	1.00	1
NODE51	2	>1000 (1.54%)	0.02	0.892
NODE53	2	13.91 (1.12%)	0.04	1
NODE56	2	0.68 (44.08%)	1.00	1
DGRICYP6G1A	1	3.12 (100.00%)	0.5	1
DGRICYP6G1B	1	0.11 (100.00%)	1.00	1
DMELCYP6G1	2	8.78 (11.98%)	0.01	0.334
DMOJCYP6G1	2	1.04 (22.53%)	0.49	1
DVIRCYP6G1A	1	0.05 (100.00%)	1.00	1
DVIRCYP6G1B	1	0.20 (100.00%)	1.00	1
DVIRCYP6G1C	1	1.18 (100.00%)	0.5	1
SFLACYP6G1A	1	1.00 (100.00%)	1.00	1

BRANCH	RATES	MAX. DN/DS	UNCORRECTED P-VALUE	CORRECTED P-VALUE
SFLACYP6G1B	1	0.29 (100.00%)	1.00	1
SFLACYP6G1C	2	26.81 (0.91%)	0.01	0.501
SFLACYP6G1D	1	0.14 (100.00%)	1.00	1
SFLACYP6G1E	1	0.18 (100.00%)	1.00	1
SFLACYP6G1F	1	0.01 (100.00%)	1.00	1
SFLACYP6G1H	1	0.67 (100.00%)	1.00	1
SGRACYP6G1A	1	0.16 (100.00%)	1.00	1
SGRACYP6G1B	1	0.25 (100.00%)	1.00	1
SGRACYP6G1C	2	1.96 (20.69%)	0.20	1
SGRACYP6G1D	1	0.28 (100.00%)	1.00	1
SHSUCYP6G1A	1	0.11 (100.00%)	1.00	1
SHSUCYP6G1B	1	0.09 (100.00%)	1.00	1
SMONCYP6G1A	1	0.09 (100.00%)	1.00	1
SMONCYP6G1B	1	0.00 (100.00%)	1.00	1
SMONCYP6G1C	1	0.00 (100.00%)	1.00	1
SMONCYP6G1F	1	0.15 (100.00%)	1.00	1
SMONCYP6G1G	1	0.25 (100.00%)	1.00	1
SMONCYP6G1H	1	0.22 (100.00%)	1.00	1
SPALCYP6G1A	1	0.05 (100.00%)	1.00	1
SPALCYP6G1C	1	0.02 (100.00%)	1.00	1
SPALCYP6G1D	1	0.00 (100.00%)	1.00	1
SPALCYP6G1B	1	0.02 (100.00%)	1.00	1

Supplementary Table 2: RNA sequencing and alignment quality

SRA accession number	Tissue type	Treatment	Number of reads (after trimming)	Percent aligned
SRR514175	whole body	Col-0 1	87.45%	13115188
SRR514175	whole body	Col-0 2	48.49%	9892004
SRR514175	whole body	GKO 1	48.71%	23139316
SRR514175	whole body	GKO 2	75.69%	18376014
SRR7649192	midgut	GKO 1	25686019	90.35%
SRR7649193	midgut	GKO 2	26644834	90.77%
SRR7649194	midgut	PAD 1	32302385	89.99%
SRR7649195	midgut	PAD 2	25657196	89.94%