Evolution in an Ancient Detoxification Pathway Is Coupled with a Transition to Herbivory in the Drosophilidae

Andrew D. Gloss,*1 Daniel G. Vassão,‡2 Alexander L. Hailey,‡3 Anna C. Nelson Dittrich,‡,1 Katharina Schramm,2 Michael Reichelt,2 Timothy J. Rast,1 Andrzej Weichsel,3 Matthew G. Cravens,1 Jonathan Gershenzon,2 William R. Montfort,3 and Noah K. Whiteman*1

1Department of Ecology and Evolutionary Biology, University of Arizona
2Department of Biochemistry, Max Planck Institute for Chemical Ecology, Jena, Germany
3Department of Chemistry and Biochemistry, University of Arizona

*These authors contributed equally to this work.

*Corresponding author: E-mail: agloss@email.arizona.edu; whiteman@email.arizona.edu.
Associate editor: Koichiro Tamura

Abstract

Chemically defended plant tissues present formidable barriers to herbivores. Although mechanisms to resist plant defenses have been identified in ancient herbivorous lineages, adaptations to overcome plant defenses during transitions to herbivory remain relatively unexplored. The fly genus Scaptomyza is nested within the genus Drosophila and includes species that feed on the living tissue of mustard plants (Brassicaceae), yet this lineage is derived from microbe-feeding ancestors. We found that mustard-feeding Scaptomyza species and microbe-feeding Drosophila melanogaster detoxify mustard oils, the primary chemical defenses in the Brassicaceae, using the widely conserved mercapturic acid pathway. This detoxification strategy differs from other specialist herbivores of mustard plants, which possess derived mechanisms to obviate mustard oil formation. To investigate whether mustard feeding is coupled with evolution in the mercapturic acid pathway, we profiled functional and molecular evolutionary changes in the enzyme glutathione S-transferase D1 (GSTD1), which catalyzes the first step of the mercapturic acid pathway and is induced by mustard defense products in Scaptomyza. GSTD1 acquired elevated activity against mustard oils in one mustard-feeding Scaptomyza species in which GSTD1 was duplicated. Structural analysis and mutagenesis revealed that substitutions at conserved residues within and near the substrate-binding cleft account for most of this increase in activity against mustard oils. Functional evolution of GSTD1 was coupled with signatures of episodic positive selection in GSTD1 after the evolution of herbivory. Overall, we found that preexisting functions of generalized detoxification systems, and their refinement by natural selection, could play a central role in the evolution of herbivory.

Key words: plant–herbivore interactions, glutathione S-transferase, detoxification, gene duplication, Drosophila, isothiocyanate.

Introduction

Herbivorous insects are remarkably successful: Approximately half of all insect species are herbivorous (Schoonhoven et al. 2005), and herbivorous lineages diversify more rapidly than their nonherbivorous relatives (Mitter et al. 1988). Paradoxically, however, evolutionary transitions to herbivory have occurred in only 9 of 30 extant insect orders (Mitter et al. 1988), despite the evolutionary success of established herbivore lineages. This suggests that feeding on living plant tissue is an evolutionary hurdle that is difficult to overcome.

Plant chemical defenses have been proposed as a major barrier to the evolution of herbivory (Strong et al. 1984; Mitter et al. 1988). Evaluating this hypothesis requires an understanding of the mechanisms enabling nascent herbivores to detoxify plant defensive compounds. Specifically, it is unclear if overcoming plant defensive toxins during evolutionary transitions to herbivory requires metabolic innovations, which may explain why transitions to herbivory are absent in many insect orders, or if newly evolved herbivores can detoxify plant toxins using generalized detoxification systems present in their nonherbivorous ancestors. Our knowledge of adaptations during transitions to herbivory is limited because most herbivorous insect species are within lineages where herbivory arose greater than 150 Ma (Labandeira 1998). Comparative genetic, metabolic, and physiological studies of younger herbivorous lineages and their nonherbivorous relatives therefore hold promise to offer insight into the adaptations facilitating survival exclusively on living plant tissues.

We investigated the mechanism to detoxify plant defensive compounds employed by recently evolved herbivores in the genus Scaptomyza (Scaptomyza flava, S. nigrita, and closely related lineages). Larvae of these species are leaf miners that consume and live within leaves of mustard plants (Brassicaceae), including the model plant Arabidopsis thaliana. Mustard-feeding Scaptomyza are nested phylogenetically within the primarily microbe-feeding Drosophila, and the
ancestor of the mustard-feeding clade transitioned to feeding on living plant tissue approximately 6–15 Ma (Whiteman et al. 2011, 2012; Lapoint et al. 2013). Scaptomyza flava feeds on a much broader range of mustard species than S. nigrita, which feeds on only one species of mustard (Wheeler 1952; Wheeler and Takada 1966; Collinge and Louda 1989; Louda and Collinge 1992; Martin 2004; Whiteman et al. 2011) (supplementary note, Supplementary Material online). Differences in the breadth of plant-derived toxins encountered by each species may result in different selective pressures on detoxification pathways and enzymes.

Although nonherbivorous drosophilids likely encounter toxins constitutively present in feeding substrates, leaf-mining larvae of S. flava and S. nigrita consume defensive compounds that accumulate to high levels in mustard leaves in response to feeding damage (Whiteman et al. 2011, 2012). The primary chemical defense compounds in mustards are derived from glucosinolates, nontoxic thioglucosides that are activated by myrosinases to form toxic isothiocyanates (ITCs or mustard oils) and other hydrolysis products upon tissue damage (Halkier and Gershenzon 2006). ITCs are highly electrophilic, and their insecticidal activity (Winde and Wittstock 2011) is thought to result primarily from disruption of proteins through conjugation to exposed nucleophilic residues (Mi et al. 2011) and cleavage of disulfide bonds (Kawakishi and Kaneko 1985). Although drosophilids utilize a tremendous diversity of feeding substrates (Markow and O’Grady 2005; Lapoint et al. 2013), glucosinolate production is narrowly restricted almost exclusively to mustards and related plant families in the order Brassicales, and glucosinolates are not known to be produced by microbes (Halkier and Gershenzon 2006). Thus, glucosinolate defenses were likely a novel dietary component associated with the transition to mustard feeding within Scaptomyza.

Interactions between herbivores and glucosinolate-bearing host plants have been a fruitful model system for understanding mechanisms of resistance to plant defensive chemicals (Winde and Wittstock 2011). However, the initial adaptations that confer resistance to glucosinolate defenses during the transition to mustard feeding—either from herbivorous or nonherbivorous ancestors—are unknown. On one hand, previously studied insects that specialize on mustards, including representatives from insect orders that span over 300 million years of divergence from a common ancestor, have evolved efficient mechanisms to sequester glucosinolates or otherwise prevent the formation of ITCs (sawfly, Athalia rosae, Muller et al. 2001; harlequin bug, Murgantia histrionica, Alabadi et al. 2002; diamondback moth, Plutella xylostella, Ratzen et al. 2002; cabbage aphid, Brevicoryne brassicae, Kazana et al. 2007; cabbage butterflies, Pieridae, Wheat et al. 2007). These mustard specialists are relatively insensitive to glucosinolate defenses (Raybould and Moyes 2001; Muller et al. 2010). Accordingly, transitions to mustard feeding might require the evolution of specific adaptations to prevent ITC formation. On the other hand, some insects with broader host ranges that occasionally feed on mustards metabolize ITCs using widely conserved, generalized detoxification mechanisms (Wadleigh and Yu 1988; Schramm et al. 2012) and remain susceptible to glucosinolate defenses (Winde and Wittstock 2011), although these studies have focused on Lepidoptera with broad host plant ranges. Generalized detoxification mechanisms might provide an initial route to resist glucosinolate defenses in insects transitioning to herbivory of mustards.

Glucosinolates slow S. flava larval development (Whiteman et al. 2011, 2012). Therefore, we hypothesized that mustard-feeding Scaptomyza would lack a mechanism to prevent ITC formation, unlike mustard specialists within more ancient herbivorous lineages. To test this hypothesis, we performed metabolic profiling to search for ITCs in larvae feeding on glucosinolate-bearing leaves. Second, we hypothesized that mustard-feeding Scaptomyza would detoxify ITCs...
using a generalized pathway present in their ancestors. We searched for ITC-derived metabolites known to be produced through generalized detoxification pathways, and for novel ITC metabolites, following glucosinolate or ITC ingestion in mustard-feeding Scaptomyza and in the nonherbivorous relative Drosophila melanogaster. We found that both nonherbivorous and mustard-specialist drosophilid flies use the widely conserved mercapturic acid pathway to detoxify mustard oils.

Given that mustard-feeding Scaptomyza detoxify ITCs using the mercapturic acid pathway, we predicted that an enzyme controlling the pathway’s substrate specificity would evolve enhanced activity against ITCs following the transition to mustard feeding. Glutathione S-transferases (GSTs) are the only enzymes known to catalyze the first step of the mercapturic acid pathway, and thus mediate the pathway’s specificity for different toxins (Habig et al. 1974). GSTs form a large multigene family in metazoans. Up to 45 genes encoding GSTs are found in single Drosophila species with sequenced genomes (Low et al. 2007). Although multiple GSTs, downstream enzymes in the mercapturic acid pathway, and toxin transport enzymes may have been targets of natural selection driven by dietary toxins during the evolution of herbivory, glutathione S-transferase D1 (GSTD1) could play a particularly important role in the detoxification of ITCs. In D. melanogaster, GSTD1 has exceptionally broad substrate specificity, and it is one of the most efficient GSTs catalyzing the conjugation of ITCs to glutathione in vitro (Saisawang et al. 2012). GSTD1 has also been implicated in putatively adaptive responses to environmental toxins in drosophilids, such as resistance to insecticide (Low et al. 2010) and to toxins present in the larval substrates of cactophilic Drosophila species (Matzkin 2008). GSTD1 expression in S. flava is induced by ingestion of jasmonate-dependent plant defenses, which include glucosinolates (Whiteman et al. 2011). Because GSTD1 evolves under strong purifying selection in Drosophila (Low et al. 2007), characterizing the function and evolution of GSTD1 in mustard-feeding Scaptomyza could offer insight into how ecologically important, but conserved, detoxification enzymes evolve during transitions to herbivory.

We conducted in vitro enzyme activity assays to test whether GSTD1 from mustard-feeding Scaptomyza detoxify ITCs more rapidly than orthologous copies from other drosophilids. Then, we solved the crystal structure of a duplicated GSTD1 enzyme from S. nigrita to identify amino acid substitutions in enzyme regions likely to affect substrate specificity, and we tested whether these residues affect the rate of ITC detoxification by GSTD1 in vitro. Finally, we tested whether GSTD1 has evolved under neutral, purifying, and/or positive selection in the mustard-feeding Scaptomyza since their divergence from a nonherbivorous Scaptomyza ancestor. Our data suggest that the transition to herbivory, and interactions with plant toxins in particular, can drive the functional evolution of generalized detoxification enzymes. However, additional studies are needed to test the hypothesis that GSTD1 facilitates resistance to mustard oils in Scaptomyza.

**Results**

**Metabolism of Glucosinolates**

Previous studies found that chewing herbivores specializing on mustards block formation of toxic ITCs from glucosinolates during feeding (Ratzka et al. 2002; Wheat et al. 2007; Winde and Wittstock 2011). To determine whether a more recently evolved mustard specialist also prevents ITC formation, we profiled glucosinolate metabolites in S. flava larvae reared on glucosinolate-deficient A. thaliana plants (Whiteman et al. 2012) infused with radiolabeled 4-methylsulfinylbutyl glucosinolate, the predominant glucosinolate in A. thaliana accession Col-0 (Brown et al. 2003). Using this unbiased approach, we detected two glucosinolate hydrolysis products in the larvae of S. flava, ITCs and nitriles, and no intact glucosinolates (fig. 2A and supplementary fig. S1A, Supplementary Material online).

In addition to free ITCs, we detected ITC-derived metabolites from the mercapturic acid pathway in the larvae of S. flava (fig. 2A and supplementary fig. S1A, Supplementary Material online). Products in the mercapturic acid pathway are formed through the conjugation of the tripeptide glutathione (GSH) to a toxin and subsequent hydrolytic modification (Boyland and Chasseaud 1969) (fig. 2B). To determine whether ITCs are detoxified using the mercapturic acid pathway across mustard-feeding Scaptomyza, we first inferred the evolutionary relationship among eight mustard-feeding Scaptomyza lineages (fig. 2C). Because the taxonomy of mustard-feeding Scaptomyza is not well resolved, we used a maximum cutoff of 3% sequence divergence at the mitochondrial genes COI and COII to delimit four major clades. We then performed a targeted search for mercapturic acid products in the larvae from one lineage from each clade. Larvae in all four lineages metabolized ITCs using the mercapturic acid pathway (fig. 2C and supplementary fig. S1B and C, Supplementary Material online). We also detected ITC-derived products arising from the mercapturic acid pathway in the adults of D. melanogaster following experimental consumption of ITCs (supplementary fig. S1D, Supplementary Material online).

**Duplication of GSTD1, a Candidate ITC Detoxification Enzyme, in Mustard-Feeding Scaptomyza**

Although GSTD1 is present as a single copy in most Drosophila species (Low et al. 2010), we discovered a novel duplication of GSTD1 in S. nigrita, and we have named these paralogs GSTD1A and GSTD1B, respectively (fig. 3). The presence of both paralogs could be verified in many S. nigrita populations in North America (supplementary table S1, Supplementary Material online) and in two of the three lineages in clade II using gene-specific polymerase chain reaction (PCR) primers and sequencing (fig. 3). However, only a single functional copy of GSTD1 was found in S. flava from Belmont, MA, through search of a draft genome assembly, and through PCR screens in other clade III and IV lineages (supplementary materials and methods, Supplementary Material online). The most parsimonious reconciliation of the gene and species trees inferred through Bayesian phylogenetic analyses requires only a single
FIG. 2. A common mechanism for ITC detoxification in mustard-feeding Scaptomyza. (A) Distribution of 4msob-glucosinolate-derived products detected in S. flava (Massachusetts) larvae feeding on leaves from A. thaliana glucosinolate-knockout mutants infused with [14C]-4msob-glucosinolate. Mercapturic acid products were the only insect-transformed metabolites detected. NAC: N-acetylcysteine. (B) Inferred route of ITC metabolism in Scaptomyza through conjugation of ITCs to GSH and subsequent transformation using the mercapturic acid pathway. (C) Left: Bayesian 50% majority-rule consensus phylogeny of mustard-feeding Scaptomyza from a partitioned alignment of six genes, rooted with S. pallida. Bayesian posterior probabilities and maximum-likelihood bootstrap values are indicated above and below each node, respectively. Four clades were designated using a maximum divergence cutoff of 3% at the genes COI and COII. Right: ITC-derived metabolites detected in the larvae from four mustard-feeding Scaptomyza lineages using HPLC-MS/MS. The chromatogram for each lineage shows overlaid MS/MS ion traces from compounds 1 to 5. Profiled lineages are indicated by dotted lines. Larvae were fed A. thaliana in the laboratory and profiled for 4-msob-glucosinolate derivatives with the exception of S. nigrita (CO), which could not be reared in the laboratory. Wild-collected S. nigrita (CO) larvae were therefore profiled for derivatives of PE glucosinolate (as well as isopropyl, isobutyl, and sec-butyl glucosinolates, see supplementary fig. S1, Supplementary Material online), which is present in their host plant. Unlike 4msob-ITC, the free ITCs formed from these glucosinolates cannot be detected using the method we employed, and thus only ITC-derived metabolites were profiled in S. nigrita larvae. Numbers above peaks correspond to compounds in panel (B). Collection localities in the United States are indicated following species/lineage names using abbreviated state names; NLD: Netherlands.
duplication event and suggests that the duplication occurred in the ancestor of clades I and II (fig. 2C). A search for GstD1 pseudogenes in the S. flava draft genome assembly does not refute this scenario: Although the assembly contained putative GstD1 pseudogenes, none nested phylogenetically with either GstD1 paralog from S. nigrita (supplementary fig. S2, Supplementary Material online).

An estimate of the timing of the duplication giving rise to GstD1A and GstD1B lends further support to the hypothesis that the duplication occurred in the ancestor of clade I and II, after the ancestors of S. flava and S. nigrita diverged. Maximum-likelihood estimates of nucleotide divergence for all pairwise comparisons of GstD1A and GstD1B sequences in our study ranged from 0.027 to 0.074 substitutions per synonymous site. Given that pairwise sequence divergence at putatively neutrally evolving sites is equal to twice the product of the mutation rate and divergence date, we estimated the age of the GstD1 duplication using a long-term mutation rate for the Hawaiian Drosophila (Tamura et al. 2004). The estimated duplication date, 1.2–3.3 Ma, is more recent than the divergence date of S. flava and S. nigrita inferred from a fossil-calibrated phylogeny (Lapoint et al. 2013). Although gene conversion events between GstD1A and GstD1B could bias this estimate toward a younger date, population genetic analyses (Betran et al. 1997) found no evidence for gene conversion among GstD1A and GstD1B in S. nigrita. A phylogenetic analysis (Pond et al. 2006) did not suggest that recombination has occurred between GstD1A

**Fig. 3.** Phylogeny of GstD1 across Scaptomyza and Drosophila. Analyses were performed using a nucleotide alignment of the full GstD1 coding sequence (630 bp) partitioned by codon position with the GTR + G substitution model. Bayesian posterior probabilities (generated using MrBayes) and maximum-likelihood bootstrap values (generated using RAxML) are shown on the Bayesian 50% majority-rule consensus tree. “GstD1A” and “GstD1B” refer to paralogous copies of GstD1 within the mustard-feeding Scaptomyza. The tree was rooted using the divergence between the subgenera Sophophora and Drosophila.
and GSTD1B since divergence of clades I and II. However, neither analysis is able to detect gene conversion events prior to the divergence of clades I and II.

**Functional Evolution of GSTD1 in Mustard-Feeding *Scaptomyza***

If GSTD1 plays a role in ITC detoxification, GSTD1 from mustard-feeding *Scaptomyza* might be expected to metabolize ITCs more efficiently than orthologs from species that do not feed on mustard plants. We compared activity profiles of heterologously expressed and purified GSTD1 protein from *S. flava* (SflaGSTD1; sequenced from the Massachusetts population), *S. nigrita* (SnigGSTD1A and SnigGSTD1B; sequenced from the Colorado population), and eight nonherbivorous *Scaptomyza* and *Drosophila* species that diverged from a common ancestor approximately 70 Ma (Whiteman et al. 2012). In in vitro assays, SnigGSTD1A conjugated seven of the eight structurally diverse ITCs to GSH more rapidly than homologous GSTD1 enzymes from other species, including the ITCs derived from the five most abundant ITC-yielding glucosinolates in the host plant of *S. nigrita* (Rodman and Louda 1985) (fig. 4 and supplementary table S2, Supplementary Material online). GSTD1A enzymes from two other mustard-feeding lineages, *S. nr. nigrita* (OR) and *S. nr. nigrita* (ID), also had higher activity against ITCs than GSTD1 from microbe-feeding *Drosophila* (supplementary fig. S3, Supplementary Material online). SflaGSTD1 exhibited high specific activity only for phenethyl ITC (PEITC), but not other ITCs (fig. 4).

Aromatic ITCs, such as benzyl ITC (BITC), are exceptionally toxic (Wadleigh and Yu 1988; Winde and Wittstock 2011). Their glucosinolate precursors are found across diverse mustard species, including in *Cardamine cordifolia*, the only known host plant of *S. nigrita* (Fahey et al. 2001). Therefore, we conducted a more extensive investigation on the efficiency of BITC conjugation to GSH by GSTD1 from *S. nigrita* and other drosophilids. We measured the activity of each enzyme across a range of BITC concentrations under a saturating GSH concentration. Two kinetic constants can be determined from this data: $k_{cat}$, the maximum BITC turnover rate per unit of enzyme, and $K_m$, the BITC concentration at which reaction velocity is half of the maximum velocity for that enzyme. In some cases, SnigGSTD1A did not have more efficient values of $k_{cat}$ or $K_m$ than GSTD1 enzymes from other species (one-sided Mann–Whitney U test, $P < 0.05$ for seven of the ten pairwise comparisons of $k_{cat}$ and eight of the ten comparisons of $K_m$ after correcting for multiple tests; fig. 4). However, the specificity constant, $k_{cat}/K_m$, which predicts the rate at which BITC is conjugated to GSH under the low BITC and high GSH concentrations likely predominating in cells

![Fig. 4. *Scaptomyza nigrita* GSTD1A is the most efficient GSTD1 ortholog at detoxifying ITCs across the Drosophilidae. Specific activity ± SEM, the velocity at which an enzyme conjugates a substrate to GSH under fixed assay conditions, was determined for eight ITCs ($N = 5–7$ replicates across 2–3 independent experiments); specific activity values and statistically significant pairwise comparisons are reported in supplementary table S2, Supplementary Material online. Cell shading reflects standard deviations of a measured value from the column mean. Michaelis–Menten kinetic constants ± SEM were determined for BITC by nonlinear regression using variable ITC concentrations at saturating GSH concentration ($N \geq 3$ independent replicates). Bolded values are not significantly different than values in the same column for SnigGSTD1A (Mann–Whitney U test). The electrophilic carbon is circled in each ITC structure. $k_{cat}$, the substrate turnover rate at $V_{max}$; $K_m$, the substrate concentration at which the reaction rate is half of $V_{max}$.](http://mbe.oxidjournals.org/)
had a lower affinity (higher $K_m$) for GSH (reflected in $K_m$) by SnigGSTD1A (supplementary fig. S4, Supplementary Material online). Our activity measurements suggest that GSTD1 could increase the rate of BITC detoxification by $2\sim3$ orders of magnitude relative to the uncollated reaction under physiologically relevant conditions, even in drosophilid species that do not naturally encounter ITCs (supplementary table S3, Supplementary Material online). However, GSTs other than GSTD1 may be important for ITC detoxification in vivo.

Activity assays using three structurally diverse non-ITC substrates of GSTs did not reveal a loss of substrate breadth or a change in affinity for GSH (reflected in $K_m$) by SnigGSTD1A (supplementary fig. S4, Supplementary Material online). On the other hand, SnigGSTD1B exhibited little activity against ITCs or these other substrates, and also had a lower affinity (higher $K_m$) for GSH (fig. 4 and supplementary fig. S4, Supplementary Material online).

Molecular Basis of Elevated Activity Against ITCs in a Paralogous GSTD1 Enzyme from Mustard-Feeding Scaptomyza

To determine how substitutions in SnigGSTD1A affect enzyme topology, we used X-ray crystallography to solve the crystal structure of SnigGSTD1A in complex with GSH (fig. 5A and B and supplementary table S4, Supplementary Material online). As is typical for GSTs, GSH was bound adjacent to a hydrophobic toxin-binding pocket (fig. 5B). Comparison of SnigGSTD1A with the structure of D. melanogaster GSTD1 (DmelGSTD1, PDB ID: 3EIN) (Low et al. 2010) revealed no dramatic changes to the structure of the protein (RMSD = 0.45 Å, Cα atoms).

SnigGSTD1A differed from orthologous GSTD1 enzymes in other drosophilids at 20–29 (of 209) amino acids. To identify candidates among these amino acid differences that may confer enhanced activity of SnigGSTD1A against ITCs, we identified amino acid substitutions in SnigGSTD1A at conserved sites in regions known to impact substrate specificity and catalytic efficiency in related Delta-class GSTs (fig 5C), described in detail below. Because the activity of GSTD1 in the ancestor of S. nigrita prior to the duplication of GSTD1 is unknown, we considered unique substitutions in SnigGSTD1A and substitutions shared by SnigGSTD1A and SnigGSTD1B as potential candidates underlying elevated activity of SnigGSTD1A against ITCs. Although resurrecting and performing site-directed mutagenesis on inferred ancestral proteins could offer insight into whether GSTD1 acquired higher activity against ITCs before or after the duplication in ancestral Scaptomyza, a marginal reconstruction approach (Yang 2007) could not infer the ancestral states at four amino acids (posterior probability ranging from 0.5 to 0.7). Therefore, we instead performed reciprocal replacements of candidate amino acids between SnigGSTD1A and DmelGSTD1—either singly or in combination with other nearby amino acids—and we measured the rate at which these modified enzymes metabolize ITCs.

Hydrophobic Binding Cleft

Changes in substrate specificity and catalytic efficiency in insect Delta-class GSTs often arise from substitutions to residues contributing to a hydrophobic toxin-binding cleft, termed the H-site (Vararattanavech et al. 2006; Wongsantichon et al. 2010). Residues forming the H-site wall (Low et al. 2010) are highly conserved in GSTD1 across drosophilids: We observed only a single substitution among 30 nonmustard-feeding species of Drosophila and Scaptomyza, with the exception of an evolutionarily labile aromatic amino acid at the periphery of the H-site. Interestingly, one amino acid substitution in the hydrophobic H-site wall was fixed in SnigGSTD1A following the GSTD1 duplication event (T166S). The same site is highly conserved across GSTD1 in schizophoran flies, including in microbe-feeding species of Drosophila and Scaptomyza (fig. 5C). Adjacent H-site residues are shifted in SnigGSTD1A relative to DmelGSTD1 (F204 and Y207 in fig. 5D and F), putatively due to the loss of a methyl group contributing to the H-site wall due to the T166S substitution. Reversing the T166S substitution lowered the specificity constant ($k_{cat}/K_m$) for BITC in SnigGSTD1A, and introducing an S166T substitution into DmelGSTD1 increased the specificity constant ($P < 0.05$ for both comparisons, one-way analysis of variance with Tukey’s HSD; fig. 5G). In both cases, specificity constants for BITC were altered by approximately 50%. A similar trend was observed for specific activity measured at high concentrations of BITC, PEITC, and an aliphatic ITC derived from glucosinolates in the host plant of S. nigrita (supplementary fig. S5, Supplementary Material online). However, these changes in activity were more subtle than the changes to the specificity constant, $k_{cat}/K_m$, suggesting that the effect of this substitution on enzyme efficiency is magnified at low, physiologically relevant ITC concentrations.

Residues Not Exposed to the Binding Cleft

Two amino acid substitutions unique to SnigGSTD1A reorganize packing of buried residues in an aromatic zipper motif near residue 166 (fig. 5C and E). Additionally, unlike most drosophilid GSTD1 enzymes, SnigGSTD1A has a polar amino acid residue at site 202, which protrudes from $\alpha$-helix 8 near the hydrophobic binding cleft (fig. 5C and F). Site 202 was identified as one of two sites putatively evolving under positive selection across Scaptomyza and the subgenus Drosophila as a whole (supplementary table S6, Supplementary Material online). After substituting all four candidate amino acids in the zipper motif, $\alpha$-helix 8, and H-site from SnigGSTD1A into DmelGSTD1, the modified DmelGSTD1 enzyme had a specificity constant ($k_{cat}/K_m$) for BITC that was greater than wild-type DmelGSTD1 and not different from SnigGSTD1A (fig. 5G).

Interface of Toxin and GSH Binding Clefts

GSTD1 enzymes across mustard-feeding Scaptomyza have undergone multiple substitutions at highly conserved amino acids (sites 33, 35–36) in a loop region between the toxin-
binding H-site and GSH binding cleft. In SnigGSTD1A, these residues form a novel hydrogen bond network absent in the structure of DmelGSTD1 (fig. 5C and D). However, exchanging amino acids at these positions between DmelGSTD1 and SnigGSTD1A—either alone (fig. 5G) or in conjunction with the other four substitutions discussed previously (supplementary table S5, Supplementary Material online)—did not affect the activity against ITCs.

Rapid Molecular Evolution of GstD1 in Mustard-Feeding Scaptomyza

To illuminate patterns of selection driving the evolution of GstD1 across drosophilids, we used phylogenetic models of evolution that rely on the rates of nonsynonymous and synonymous substitutions expected under different selection pressures. Episodic positive selection results in an excess of normalized nonsynonymous substitutions compared with synonymous substitutions ($\omega > 1$), whereas synonymous substitutions are fixed much more frequently than nonsynonymous substitutions under purifying selection ($\omega < 1$).

Across the genus Scaptomyza and the subgenus Drosophila, in which Scaptomyza and the Hawaiian Drosophila are nested phylogenetically, we found a strong signal of gene-wide purifying selection in GstD1 (one-ratio model, gene-wide $\omega = 0.1100$). A model allowing some codons to evolve under positive selection was a better fit than a model constraining all codons to evolve under
purifying selection or neutral evolution (models M7 versus M8, likelihood-ratio test, \( P = 0.0066 \); supplementary table S6, Supplementary Material online). Although identification of single codons under positive selection should be interpreted cautiously, there was strong support (posterior probability \( > 0.95 \)) that codons at position 133 and 202 of GstD1 were evolving under positive selection across drosophilids. However, the model allowing positive selection was not a significantly better fit than a conservative null model less prone to false positives when some codons are evolving neutrally (models M8a vs. M8, \( P > 0.1 \); supplementary table S6, Supplementary Material online). Thus, conservative model comparisons did not find strong evidence for episodic positive selection (\( \omega > 1 \)) acting on specific codons in GstD1 across the subgenus Drosophila and genus Scaptomyza generally.

Next, we tested for an elevated rate of nonsynonymous substitutions in GstD1 in individual lineages of mustard-feeding Scaptomyza (Yang 1998). Although evolution of GstD1 was dominated by purifying selection across microbe-feeding drosophilids, \( \omega \) was elevated after the evolution of herbivory in Scaptomyza (branches indicated as -A,B; -A; -B; and -F in fig. 6). Interestingly, however, \( \omega \) along more recent branches of the phylogeny for SnigGstD1A, SnigGstD1B, or SflaGstD1 was not statistically different than in Drosophila and Scaptomyza as a whole. This is striking because the presence of locally adaptive or slightly deleterious nonsynonymous polymorphism, which contributes proportionally less than synonymous polymorphism to among-species divergence, would be expected to inflate \( \omega \) in terminal branches for closely related lineages relative to deeper branches of the phylogeny (Peterson and Masel 2009). Although \( \omega \) was elevated in GstD1 following the evolution of herbivory, only two branches showed \( \omega > 1 \), and the models allowing \( \omega > 1 \) along either branch did not fit significantly better than a model fixing \( \omega = 1 \) (\( \ln L = -4009.958 \); for -A, \( \ln L = -4013.494 \) for -A,B; likelihood-ratio test \( P > 0.05 \) in both cases). More powerful tests to detect a subset of codons evolving under positive selection (\( \omega > 1 \)) (Zhang et al. 2005) could not be applied to individual branches because these tests may be prone to false positives when the total number of substitutions is small (Nozawa et al. 2009).

Therefore, although elevated \( \omega \) in mustard-feeding Scaptomyza could result from episodic positive selection affecting only a subset of codons in GstD1, phylogenetic models of variation in \( \omega \) cannot rule out past relaxation of purifying selection as the cause of this pattern.

The McDonald–Kreitman test detects an excess of nonsynonymous divergence relative to expectations based on current patterns of polymorphism. Polarized McDonald–Kreitman tests revealed an excess of non-synonymous substitutions in SflaGstD1, SnigGstD1A, and SnigGstD1B (table 1 and supplementary table S9, Supplementary Material online). The proportion of amino acid substitutions predicted to have been fixed by positive selection (\( \omega \); Smith and Eyre-Walker 2002) exceeded 85% for each gene (table 1), although this statistic fluctuates stochastically and should be interpreted cautiously (Nei et al. 2010). Genes without roles in detoxification did not yield statistically significant McDonald–Kreitman tests or show elevated \( \omega \) in S. flavia or S. nigrita (supplementary tables S7 and S8, Supplementary Material online). This suggests that evidence for positive selection on SnigGST1DA, SnigGST1DB, and SflaGST1D was not due to demographic processes, such as a population bottleneck, which can leave genome-wide molecular signatures resembling positive selection. Although statistical tests for positive selection may yield false inferences when assumptions are violated, the results of our independent tests are consistent with a scenario in which the transition to herbivory in Scaptomyza is associated with positive selection driving evolution of GstD1.

### Supplementary Material

For supplementary material, please refer to the journal's supplementary material section.
**Table 1.** Polarized McDonald–Kreitman Tests for GstD1 in Mustard-Feeding Scaptomyza.

<table>
<thead>
<tr>
<th>Species</th>
<th>Gene</th>
<th>Method</th>
<th>Nchr*</th>
<th>Ps*</th>
<th>Ds*</th>
<th>Dn*</th>
<th>α*</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. flava</td>
<td>GstD1</td>
<td>1</td>
<td>20</td>
<td>8</td>
<td>7</td>
<td>11</td>
<td>1</td>
<td>0.0074*</td>
</tr>
<tr>
<td>S. nigrita</td>
<td>GstD1A</td>
<td>1</td>
<td>48</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>1.0003**</td>
</tr>
<tr>
<td>S. nigrita</td>
<td>GstD1B</td>
<td>1</td>
<td>36</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>0.0427*</td>
</tr>
<tr>
<td>S. flava</td>
<td>GstD1</td>
<td>2</td>
<td>20</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>8</td>
<td>0.0177*</td>
</tr>
<tr>
<td>S. nigrita</td>
<td>GstD1A</td>
<td>2</td>
<td>48</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0.0048*</td>
</tr>
<tr>
<td>S. nigrita</td>
<td>GstD1B</td>
<td>2</td>
<td>36</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>0.0300*</td>
</tr>
</tbody>
</table>

*Method 1 included all sites (full 209 codons for S. nigrita; 182 codon partial gene sequence for S. flava). Method 2 excluded sites for which the codon identity in the ancestor of S. flava and S. nigrita (for S. flava) or S. nigrita GstD1A and GstD1B (for both S. nigrita paralogs) could not be inferred with >0.90 probability using maximum-likelihood reconstruction (see supplementary materials and methods, Supplementary Material online).

*Number of chromosomes sampled.

**Number of synonymous polymorphisms within a lineage.

*Number of nonsynonymous polymorphisms within a lineage.

*Number of synonymous substitutions within a lineage.

*Estimated proportion of nonsynonymous substitutions fixed by positive selection (ν = 1 − Dn/Pn/DsPs).

*P value from Fisher's exact test. P values are presented without correction for multiple tests. Significance after correcting for multiple tests using the method of Benjamini and Hochberg is indicated.

**Discussion**

An Ancient Metabolic Pathway Detoxifies Plant-Derived Chemicals in a Recently Evolved Herbivorous Insect Lineage

Most mustard-specialist herbivores are highly resistant to glucosinolate defenses (Raybould and Moyes 2001; Muller et al. 2010). Notably, chewing insect herbivores that specialize on mustard plants have evolved mechanisms to prevent the activation of glucosinolates into toxic ITCs (Muller et al. 2001; Ratzka et al. 2002; Wheat et al. 2007; Winde and Wittstock 2011), which otherwise occurs in mustards when leaf tissue is damaged during feeding. However, studies have focused on mustard specialist species within lineages for which the timing of the transition to herbivory is ancient or unknown. Thus, it is unclear whether highly efficient mechanisms to prevent ITC formation are necessary for evolutionary transitions to herbivory of mustards.

ITCs are highly toxic to D. melanogaster (Lichtenstein et al. 1962), a nonherbivorous relative of Scaptomyza, and may have presented a barrier to the evolution of mustard feeding. The evolutionary transition to mustard herbivory in Scaptomyza occurred without novel, highly efficient mechanisms to prevent ITC formation. Instead, lineages spanning the diversity of mustard-feeding Scaptomyza, in which herbivory evolved relatively recently (Whiteman et al. 2012; Lapoint et al. 2013), detoxify ITCs through the mercapturic acid pathway. This is also the primary mode of ITC metabolism in generalist feeders, including humans (Mennicke et al. 1988), some lepidopteran larvae (Schramm et al. 2012), and mollusks (Falk et al. 2013). In an unbiased quantitative analysis, no ITC-derived metabolites other than mercapturic acid products could be detected as metabolites produced by S. flava larvae. We found that D. melanogaster, a nonherbivorous relative of mustard-feeding Scaptomyza, also detoxifies ITCs using the mercapturic acid pathway. The ability to detoxify ITCs through the mercapturic acid pathway was therefore likely present in the earliest ancestral drosophilids and predated the evolution of mustard herbivory in Scaptomyza.

Scaptomyza flava feeds and develops more slowly when encountering glucosinolate defenses (Whiteman et al. 2012), unlike mustard specialists that prevent ITC formation (Raybould and Moyes 2001; Muller et al. 2010). Two aspects of the glucosinolate and ITC metabolism in S. flava may underlie this observation. First, detoxification of ITCs through the mercapturic acid pathway in S. flava may be nutritionally costly. The GST-catalyzed step in the mercapturic acid pathway involves the addition of a nitrogen-containing compound (GSH) to ITCs, and herbivores are thought to be nitrogen limited (Mattson 1980). Furthermore, GSH synthesis requires cysteine, which insects can obtain directly from the diet or synthesize from methionine. Cysteine and methionine are among the least abundant amino acids in plant leaves and may be growth limiting for chewing herbivores, and a large fraction (~20%) of total cysteine content is used to produce GSH (Barbehenn, Kochmanski, et al. 2013; Barbehenn, Niewiadomski, et al. 2013). Second, ITCs have direct chemotoxic effects. Free ITCs bind to nucleophilic residues of cellular proteins (Mi et al. 2011; Zhang 2012), disrupt disulfide bridges important for protein structure (Kawakishi and Kaneko 1985), and can induce cell death and apoptosis (Brown and Hampton 2011). These chemotoxic effects of ITCs likely contribute to the induction of stress response genes following ingestion of dietary glucosinolates in S. flava larvae (Whiteman et al. 2012).

Evolution of an Increased Rate of ITC Detoxification in a GST from a Mustard-Feeding Drosophilid

GSTs, a family of enzymes that catalyze the mercapturic acid pathway, have broad substrate specificity (Hayes et al. 2005). However, if ITCs imposed strong selection on GSTs following the transition to mustard herbivory in Scaptomyza, these enzymes could evolve increased activity against ITCs. We
tested this hypothesis by comparing activity against ITCs by the enzyme GSTD1, an abundant and ecologically important detoxification enzyme in drosophilids. We found that SnigGSTD1A, one of the two copies of GSTD1 in S. nigrita resulting from a gene duplication event in the mustard-feeding Scaptomyza lineage, initiates the detoxification of diverse ITCs at a rate unmatched by orthologous GSTD1 enzymes we profiled in other Drosophila. However, more radical kinetic changes have been inferred in other detoxification enzymes—such as cytochrome P450s—following host shifts or exposure to pesticides (Li et al. 2003, 2007; Joußen et al. 2012). Larger activity increases often involve enzymes that undergo rapid duplication and loss (Li et al. 2003) or form novel chimeric enzymes through recombination (Joußen et al. 2012). Purifying selection—which we found to be pervasive for GstD1 across microbe-feeding species in the genus Scaptomyza and subgenus Drosophila—might prevent similarly large changes in GSTD1 function. Furthermore, radical increases in activity against ITCs may be difficult to achieve. The specificity constants ($k_{cat}/K_m$) of drosophilid GSTD1 enzymes against two ITCs are within an order of magnitude of the largest known for these substrates among any GST (Kolm et al. 1995; Wiktelius and Stenberg 2007), and random engineering of a human GST with high activity against BITC and PEITC resulted in no more than 2-fold increases in activity (Runarsdottir and Mannervik 2010). On the other hand, because Drosophila GSTs other than GSTD1 are also able to metabolize ITCs (Saisawang et al. 2012), the absence of a larger increase in activity against ITCs may reflect a limited role for GSTD1 in ITC detoxification. Nonetheless, kinetic changes similar in magnitude to those observed in SnigGSTD1A can have large effects on organismal fitness (Walkiewicz et al. 2012), and functional evolution of SnigGSTD1A is consistent with an adaptive evolutionary response to dietary ITCs.

Although molecular evolutionary analyses suggested that GstD1 evolved under positive selection in the ancestor of S. flava, we did not observe an increase in GSTD1 activity against most ITCs in this species. SfGaGSTD1 did, however, exhibit high specific activity and catalytic efficiency ($k_{cat}/K_m$) against PEITC. Phenethyl glucosinate (which yields PEITC) and hydroxylated phenethyl glucosinolate dominate the glucosinolate profile of B. vulgaris (van Leur et al. 2006), the host plant from which the S. flava individuals used in this study were collected in Belmont, MA. Thus, a role for ITCs in driving evolution of SfGaGSTD1 cannot be rejected.

Structural Basis of Changes in GSTD1 Function Following Gene Duplication

Evolutionary genetic analyses suggest that many amino acid substitutions in SnigGSTD1A, which acquired elevated activity against ITCs, were likely fixed by positive selection following gene duplication, although forces other than glucosinolate defenses may have been responsible. Therefore, linking structural and functional changes in SnigGSTD1A could reveal new insight into adaptive structural changes during the divergence of paralogous GSTs coupled with exposure to novel selective pressures. To identify residues in SnigGSTD1A that enhanced activity against ITCs, we assayed enzyme activity after swapping candidate amino acid residues between SnigGSTD1A and DmelGSTD1, which has a specificity constant ($k_{cat}/K_m$) for BITC that is less than 40% of the specificity constant of SnigGSTD1A.

Only one substitution occurred in the hydrophobic toxin-binding pocket of SnigGSTD1A (T166S), and site-directed mutagenesis revealed that this substitution accounts for approximately 50% of the difference in specificity for ITCs in SnigGSTD1A versus DmelGSTD1. This suggests that elevated activity against ITCs by SnigGSTD1A evolved, at least partly, following gene duplication. Although the substituted site is highly conserved in drosophilid GSTD1 enzymes, a parallel substitution occurred in a GstD1 paralog resulting from gene duplication in D. repleta group, which contains cactophilic (though nonherbivorous sensu stricto) species that are also tightly linked to their host plants. The substituted (Ser) and ancestral (Thr) amino acids do not differ in charge and are of a similar size. Despite this, substitutions between Thr166 and Ser164 have independently evolved at a positively selected site in the binding pockets of mammalian Mu-class GSTs, and these substitutions alter substrate specificity (Norgård et al. 2006). Our results are consistent with previous findings that nonradical substitutions in the hydrophobic toxin-binding pocket can drive functional divergence of GSTs.

We also observed a cluster of substitutions unique to SnigGSTD1A or shared by mustard-feeding Scaptomyza species at the interface of the hydrophobic and GSH binding pockets, which form a novel hydrogen bond network. This region affects substrate specificity in another dipteran Delta-class GST (Vararattanavech et al. 2006). Furthermore, specialization for metabolism of dietary furanocoumarins by papilionid cytochrome P450 enzymes appears to result, in part, from novel stabilizing networks in the catalytic pocket that reduce flexibility and maintain an enzymatic conformation favorable for furanocoumarin metabolism (Li et al. 2003, 2004). However, this region did not affect specificity for ITCs in GSTD1. Given that this region differs between SnigGSTD1A, SnigGSTD1B, and SfGaGSTD1—one of which have likely undergone episodic positive selection—its functional importance in these enzymes warrants further investigation.

Because substitutions of exposed amino acids in the binding cleft did not fully explain differences in activity against ITCs between SnigGSTD1A and DmelGSTD1, we investigated substitutions unique to SnigGSTD1A or both SnigGSTD1A paralogs that occurred outside the binding cleft. Two substitutions rearranged packing of an aromatic zipper buried near the hydrophobic binding cleft, which affects flexibility, and in turn substrate specificity, in another dipteran Delta-class GST (Wongsantichon et al. 2010). Additionally, a charged amino acid substitution formed a novel hydrogen bond with the backbone of the mobile $\alpha$-helix, which partially occludes the hydrophobic binding cleft (Low et al. 2010). Although this may affect the stability of the $\alpha$-helix (Scholtz and Baldwin 1992), its functional importance is not clear. Swapping these three candidate substitutions in the $\alpha$-helix and
hydrophobic zipper between DmelGSTD1 and SnigGSTD1A, in conjunction with the T166S substitution to the wall of hydrophobic binding cleft, nearly completely reversed specificity constants of SnigGSTD1A and DmelGSTD1 for ITCs. The direction of these changes in activity suggests that one or more of these substitutions may have been adaptive in S. nigrita.

Although we observed only very small topological changes in regions affecting SnigGSTD1A activity against ITCs, GSTs undergo conformational shifts upon substrate binding (Wongsantichon et al. 2012), and our attempts to cocrystallize SnigGSTD1A with ITCs were unsuccessful. Therefore, our structural analysis may underestimate the extent of topological changes relevant to ITC detoxification.

A Potential Role for Relaxation of Functional Constraint on GstD1 Following a Lineage-Specific Duplication in S. nigrita

Strong purifying selection we and others (Low et al. 2007) inferred for GstD1 in drosophilids may result from pressure to maintain both broad substrate specificity and an essential developmental function. GSTD1 exhibits the broadest known substrate specificity among GSTs in D. melanogaster (Saisawang et al. 2012), and we found that the ability to metabolize a broad range of toxins is maintained across the drosophilid radiation. Furthermore, RNAi-mediated knockdown of GstD1 in D. melanogaster results in pupal-stage lethality, suggesting that GSTD1 also plays a developmental role (RNAi stock 10045R-4, National Institute of Genetics, Japan). The retention of at least one GstD1 copy across the Drosophila and Scaptomyza radiations, despite the dynamic turnover of many GSTs observed in fully sequenced Drosophila genomes (Low et al. 2007), further suggests that one or more functions performed by GSTD1 are essential in drosophilids.

When a single gene performs multiple functions, pleiotropy may constrain the simultaneous optimization of both functions, termed adaptive conflict (Piatigorsky and Wistow 1991; Des Marais and Rausher 2008). This model is particularly relevant to promiscuous detoxification enzymes that metabolize a diversity of toxic substrates, such as GSTD1 (Storz 2008). Gene duplication allows escape from adaptive conflict. Different functions are then refined in each paralog via positive selection. Both GstD1 paralogs in S. nigrita show signatures of episodic positive selection, and phylogenetic models revealed that nonsynonymous substitutions were fixed at an accelerated rate in both paralogs following gene duplication but not in the more recent history of either paralog. This pattern would be expected if different ancestral functions were refined in each paralog following duplication. Interestingly, however, SnigGSTD1B had minimal activity against all ITC and non-ITC substrates. One evolutionary scenario is that refinement of a developmental role by SnigGSTD1B enabled SnigGSTD1A to acquire increased activity against ITCs without negative pleiotropic effects. Similarly, the loss of high activity against many substrates observed in SnigGSTD1B may have been enabled by the retention of these activities in SnigGSTD1A. Characterizing differences in the timing and location of SnigGSTD1A and SnigGSTD1B expression is likely to give further insight into their functional divergence. Future studies could test for roles of both enzymes in generalized detoxification, mustard oil resistance, and development by expressing either SnigGSTD1A or SnigGSTD1B in D. melanogaster.

Differences in host breadth between S. flav a and S. nigrita may also influence the strength of purifying selection on GstD1. Although S. flav a feeds on >40 mustard species and a few unrelated host plant species (Whiteman et al. 2011), S. nigrita specializes on a single mustard species (Collinge and Louda 1989) (supplementary note, Supplementary Material online) and is likely to be exposed to a more restricted set of plant-derived toxins. The narrow host breadth of S. nigrita could have permitted the evolution of increased efficiency and specificity for ITCs at the expense of activity against other toxins not present in its only known host plant species. In an analogous situation, narrow dietary specialization may have enabled functional specialization for furanocoumarins by promiscuous cytochrome P450 enzymes in papilionid larvae (Li et al. 2003, 2004; Wen et al. 2006).

Generalized Detoxification Systems May Facilitate the Evolution of Herbivory

The larvae of most Scaptomyza species feed on microbes living in rotting vegetation (Lapoint et al. 2013). We have successfully reared S. pal lida and S. apicata, which last shared a common ancestor with herbivorous Scaptomyza >15 Ma (Whiteman et al. 2012; Lapoint et al. 2013), in rotting tissue from mustard plants (supplementary note, Supplementary Material online). The ability to detoxify mustard oils using the mercapturic acid pathway—which is shared by herbivorous mustard-feeding Scaptomyza and microbe-feeding D. mel anogaster—could have facilitated feeding on rotting and then living mustard tissues.

Interestingly, herbivory has evolved more times in flies (Diptera) than in other insect orders, including multiple times in the Drosophila radiation (Mitter et al. 1988; Whiteman et al. 2012). Comparative functional and genomic studies of these species and their closest relatives, which are not herbivores, will shed further light on the role of GSTs and other detoxification enzymes in facilitating transitions that are among the most important in the history of terrestrial life.

Conclusion

Rapid species radiations follow evolutionary transitions to herbivory, such that herbivorous insects account for approximately half of insect species diversity. However, these transitions are absent from most insect orders. Understanding the behavioral, morphological, and physiological underpinnings of the transition to herbivory is important to addressing this paradox. From a physiological perspective, plant toxins present in living plant tissue have long been hypothesized to render living plant tissues recalcitrant to invasion by insects. Although specialized herbivorous species possess derived and efficient detoxification mechanisms, it is unclear whether
more generalized detoxification mechanisms play an important role in the transition to herbivory. Here, we tested the hypothesis that generalized detoxification mechanisms may play important roles in evolutionary transitions to herbivory. We found that a recently derived lineage of drosophilid herbivore uses generalized, ancestral detoxification pathway to metabolize ITCs, a major defensive toxin in their host plants. After the transition to herbivory, an otherwise highly conserved GST, catalyzing the initiation of this pathway, may have evolved under positive selection in the herbivorous lineage. Signatures of positive selection are coupled with the evolution of higher activity against ITCs in one paralog following gene duplication. We found that a single amino acid substitution in this enzyme is critical for the increased specificity against ITCs. Given these results, and the fact that GSTs arose before the divergence of plants and animals (Sheehan et al. 2001), the generalized detoxification pathway catalyzed by GSTs could have helped to facilitate the very first transitions to herbivory in insects >350 Ma (Labandeira 1998). We anticipate that studies in other lineages will provide insight into the extent to which recently derived herbivores use generalized detoxification pathways, rather than entirely novel mechanisms, to metabolize dietary plant toxins. This could help illuminate the evolutionary origins of herbivory, one of the most widespread, complex, and ecologically important phenotypes known in metazoans.

Materials and Methods
Detailed Materials and Methods with references are described in the Supporting Information (Supplementary Material online).

Analysis of Glucosinolate and 4msob-ITC Metabolism
To enable quantitative analysis in larval S. flava, [14C]-4msob-glucosinolate was prepared biosynthetically from [U-14C]-Met, HPLC purified, and introduced into detached leaves of A. thaliana (cyp79b2 cyp79b3 myb28 myb29 in the Col-0 background) (Schramm et al. 2012). After feeding, S. flava larvae were homogenized, extracts were fractionated by HPLC, and metabolites were quantified by scintillation counting (Schramm et al. 2012). Metabolites in other species were semiquantitative. Mustard-feeding Scaptomyza larvae were fed A. thaliana in the laboratory or wild-collected from mustard plants (see supplementary materials and methods, Supplementary Material online). Adult D. melanogaster were fed with 1 mM 4msob-ITC in 5% sucrose. Compounds were analyzed by HPLC-mass spectrometry (MS)/MS and identified based on comparison of retention times to previously isolated compounds or synthetic standards (Schramm et al. 2012).

PCR and Sequencing of GstD1
Degenerate PCR primers (supplementary table S10, Supplementary Material online) were used to amplify the GstD1 coding sequences (CDS) (630 bp) from S. flava, S. nigrita, additional mustard-feeding lineages that could not be confidently identified as described species, and 16 species of Scaptomyza and Hawaiian Drosophila. Searches for new paralogs in S. flava and S. nigrita used high-efficiency thermal asymmetric interlaced PCR (Liu and Chen 2007) and, for S. flava, a preliminary genome assembly. Specific primers were designed to amplify GstD1 from individuals from single populations of S. flava and S. nigrita for population genetic analysis. Sequencing of purified PCR products or plasmids generated overlapping reads across the CDS. DNA sequences have been deposited in GenBank under accession codes listed in supplementary table S1, Supplementary Material online.

Construction of GSTD1 Expression Vectors and Enzyme Purification
The coding sequence of GstD1 (or, for S. nigrita, GstD1A and GstD1B) was amplified from each of the ten Drosophila and Scaptomyza species and inserted into the pT7-MAT-Tag-2 expression vector using restriction digest and ligation. Constructs were sequence verified before and after transformation into Rosetta 2 E. coli. Enzymes were overexpressed and purified using affinity for Glutathione Sepharose (supplementary fig. S6, Supplementary Material online).

Specific Activity and Kinetic Assays
Initial reaction rates of GSH:substrate conjugation were determined using spectrophotometry (supplementary table S11, Supplementary Material online), with published or newly determined extinction coefficients (supplementary materials and methods, Supplementary Material online). Specific activities reported are the mean of four to eight independent replicates. For kinetic assays, initial reaction rates were measured (N = 3–7) at seven to nine concentrations of the rate-limiting substrate, and Michaelis–Menten kinetic parameters were determined using nonlinear regression.

Structure Determination and Refinement
Successful crystalization conditions for SnigGSTD1A with a GSH ligand were identified by robotic screening trials and optimized using a sitting-drop vapor diffusion method. The structure was determined from X-ray diffraction data by molecular replacement using D. melanogaster GSTD1 (PDB ID: 3EIN) (Low et al. 2010), manually rebuilt, and refined. Coordinates and structure factors for the final model have been deposited in the Protein Data Bank under accession code 4I97.

Phylogenetic and Molecular Evolutionary Analysis
Phylogenetic analyses of GstD1 in drosophilids and a multigene phylogeny for mustard-feeding Scaptomyza were implemented under maximum likelihood (Stamatakis 2006) and Bayesian (Ronquist and Huelsenbeck 2003) frameworks with the general time reversible (GTR)+G substitution model. The tree topology for GstD1 in mustard-feeding Scaptomyza inferred through Bayesian analyses, presented in figure 3, was used for phylogenetic analysis of codon evolution rates (ω). Because many internal nodes in the GstD1 genealogy could not be confidently inferred, previously described species

13
relationships were used to specify the topology for nonherbivorous taxa in this analysis. Models implemented in PAML 4.5 (Yang 2007) tested for positive selection across the genus Scaptomyza and subgenus Drosophila (including the Hawaiian Drosophila) generally, and for accelerated evolution and positive selection in specific mustard-feeding lineages. McDonald–Kreitman tests (McDonald and Kreitman 1991), polarized using marginal ancestral reconstruction conducted using PAML 4.5 (Yang 2007), were also used to test for lineage-specific positive selection in S. flavus and S. nigrita. Similar analyses were conducted for additional genes to control for demographic effects.

Supplementary Material
Supplementary note, materials and methods, figures S1–S6, and tables S1–S11 are available at Molecular Biology and Evolution online (http://mbe.oxfordjournals.org/).

Acknowledgments
The authors thank the editor and two anonymous reviewers for helpful comments on the manuscript; Richard Lapoint for collection and identification of many fly specimens used in this study; Noel Kitchen for assistance with experiments; and Vahe Bandarian, Mark Beilstein, Matthew Cordes, and Megan McEvoy for use of equipment at the University of Arizona. Specimen collection permits were provided by the United States Forest Service. This work was supported by the National Institutes of Health (R01 HL062969 to W.R.M.; Shared Instrumentation Grant S10 RR025485 to A.W.; PERT fellowship 5K12GM000708-13 to A.N.D.), the National Science Foundation (DEB-1256758 to N.K.W.; DEB-1405966, DGE-1143953, and University of Arizona IGERT Genomics Fellowship DGE-0654435 to A.D.G.), the John Templeton Foundation (Grant ID #41855 to N.K.W.), the National Geographic Society (9097-12 to N.K.W.), the University of Arizona (Faculty Seed Grant, Center for Insect Science Seed Grant, and laboratory set-up grant to N.K.W.), the Rocky Mountain Biological Laboratory (fellowships to A.D.G. and N.K.W.), and the Max Planck Society (funds to J.G.). Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory, and an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the Department of Energy Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393), and the National Center for Research Resources (P41RR001209).

References


