

# The genetic basis of a plant–insect coevolutionary key innovation

Christopher W. Wheat<sup>\*†‡</sup>, Heiko Vogel<sup>\*</sup>, Ute Wittstock<sup>\*§</sup>, Michael F. Braby<sup>¶||</sup>, Dessie Underwood<sup>\*\*</sup>, and Thomas Mitchell-Olds<sup>\*††</sup>

<sup>\*</sup>Max Planck Institute for Chemical Ecology, Beutenberg Campus, Hans Knoell Strasse 8, 07745 Jena, Germany; <sup>†</sup>School of Botany and Zoology, Australian National University, Canberra ACT 0200, Australia; <sup>\*\*</sup>California State University, 1250 Bellflower Boulevard, Long Beach, CA 90840; and <sup>††</sup>Department of Biology, Duke University, Durham, NC 27708

Edited by May R. Berenbaum, University of Illinois at Urbana–Champaign, Urbana, IL, and approved October 23, 2007 (received for review July 5, 2007)

Ehrlich and Raven formally introduced the concept of stepwise coevolution using butterfly and angiosperm interactions in an attempt to account for the impressive biological diversity of these groups. However, many biologists currently envision butterflies evolving 50 to 30 million years (Myr) after the major angiosperm radiation and thus reject coevolutionary origins of butterfly biodiversity. The unresolved central tenet of Ehrlich and Raven's theory is that evolution of plant chemical defenses is followed closely by biochemical adaptation in insect herbivores, and that newly evolved detoxification mechanisms result in adaptive radiation of herbivore lineages. Using one of their original butterfly–host plant systems, the Pieridae, we identify a pierid glucosinolate detoxification mechanism, nitrile-specifier protein (NSP), as a key innovation. Larval NSP activity matches the distribution of glucosinolate in their host plants. Moreover, by using five different temporal estimates, NSP seems to have evolved shortly after the evolution of the host plant group (Brassicales) ( $\approx 10$  Myr). An adaptive radiation of these glucosinolate-feeding Pierinae followed, resulting in significantly elevated species numbers compared with related clades. Mechanistic understanding in its proper historical context documents more ancient and dynamic plant–insect interactions than previously envisioned. Moreover, these mechanistic insights provide the tools for detailed molecular studies of coevolution from both the plant and insect perspectives.

adaptive radiation | Brassicales | Pieridae | diversification | Bayesian relaxed molecular clock

The relative timing of adaptive radiations in host plants and their butterfly herbivores is controversial. Although the major angiosperm radiation occurred  $\approx 140$  to 100 million years ago (Mya), fossil data suggest that diversification of “primitive” Lepidoptera occurred before this time and butterflies radiated long after these host plants ( $< 75$  Mya) (1–3). Many espouse this recent butterfly origin, which necessarily implies a very limited role, if any, for coevolution in butterfly diversification (1–4). However, others posit a much older age of butterflies ( $> 100$  Mya), with speciation influenced by angiosperm evolution and the breakup of the supercontinent Gondwana (5, 6). This lack of consensus on both the timing of butterfly diversification, which resulted in the  $\approx 17,000$  extant species today, and the role of coevolution arises from the notably poor fossil record of Lepidoptera (5, 7). Consequently, whereas some have described likely scenarios, no studies have tested for the effects of key innovations on butterfly diversification even though coevolution through key innovations was first introduced to science using butterflies and their angiosperm host plants as exemplars (8–12). To explore the potential role of coevolution in butterfly diversification, we focus on the family Pieridae, composed of the commonly known white and sulfur butterflies. Recent significant advances in functional genomics and phylogenetics in this family provide a unique opportunity to resolve the controversies briefly reviewed above (see also ref. 5).

Pieridae use three major host plant groups: the Fabales (Legumes), the Brassicales (glucosinolate-containing plants ex-

emplified by the cabbages and *Arabidopsis*), and mistletoes. Phylogenetic reconstruction of almost 90% of the Pieridae genera (74 recognized genera plus six subgenera, based on 1,066 bp of the EF-1 $\alpha$  gene) was recently completed (13). These results indicate that Fabales feeding is the ancestral state of Pieridae (Fig. 1). The Fabales feeders are the Dismorphiinae and nearly all Coliadinae, whereas the sister to the Coliadinae, the Pierinae, primarily feed on Brassicales (Fig. 1) (12). Within Pierinae, there are two subsequent derived shifts away from glucosinolate feeding onto mistletoes and other species. Thus, the Pierinae represent a single origin of glucosinolate feeding (Fig. 1).

The evolutionary appearance of the plant order Brassicales (Eurosid II, Dicotyledons) presented a radical new chemical challenge for insect herbivores, known as the glucosinolate-myrosinase system (14). All Brassicales have this system, which is one of the best and most widely studied chemical plant defenses (15–17). Its effectiveness as an anti-herbivore defense becomes apparent upon tissue damage, such as insect feeding. Tissue damage brings the formerly compartmentalized myrosinase enzyme into contact with nontoxic glucosinolates, which it hydrolyzes into breakdown products such as isothiocyanates (18–20). Whereas *Homo sapiens* may find these breakdown products enjoyable condiments (e.g., mustard, wasabi), they are well known to be highly toxic to many insect herbivores (15, 21–23).

We have identified two independent lepidopteran detoxification mechanisms for the glucosinolate-myrosinase defense system at biochemical and molecular levels by means of functional genomics approaches, beginning with glucosinolate sulfatase (GSS) in the diamondback moth *Plutella xylostella* (Plutellidae) (24). GSS desulfates glucosinolates, producing metabolites that no longer act as substrates for myrosinases. The second, called nitrile-specifier protein (NSP), has recently been identified for the pierid butterfly *Pieris rapae* (20). NSP, expressed solely in the larval midgut, promotes the formation of nitrile breakdown products instead of toxic isothiocyanates upon myrosinase-catalyzed glucosinolate hydrolysis (20). The GSS and NSP detoxification mechanisms are distinctly different from each other, as well as the other identified host plant detoxification

Author contributions: C.W.W. and T.M.-O. designed research; C.W.W., H.V., and U.W. performed research; U.W., M.F.B., D.U., and T.M.-O. contributed new reagents/analytic tools; C.W.W., H.V., and U.W. analyzed data; and C.W.W., H.V., and U.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>†</sup>Present Address: Pennsylvania State University, Department of Biology, 208 Mueller Laboratories, University Park, PA 16802.

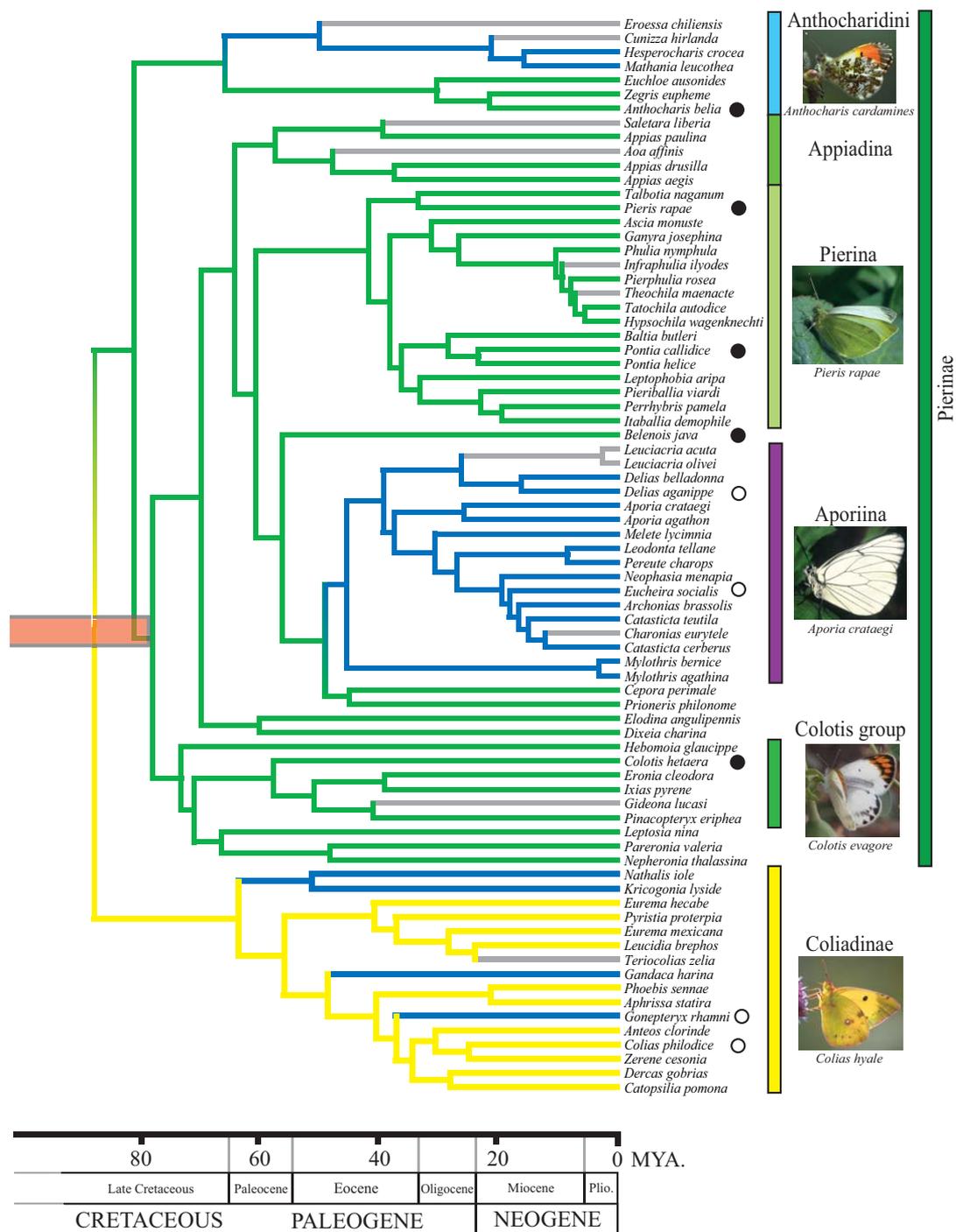
<sup>‡</sup>To whom correspondence should be addressed. E-mail: cww10@psu.edu.

<sup>§</sup>Present Address: Institute of Pharmaceutical Biology, Braunschweig University of Technology, 38106 Braunschweig, Germany.

<sup>||</sup>Present Address: Biodiversity Conservation Division, Department of Natural Resources, Environment and the Arts, P.O. Box 496, Palmerston NT 0831, Australia.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0706229104/DC1](http://www.pnas.org/cgi/content/full/0706229104/DC1).

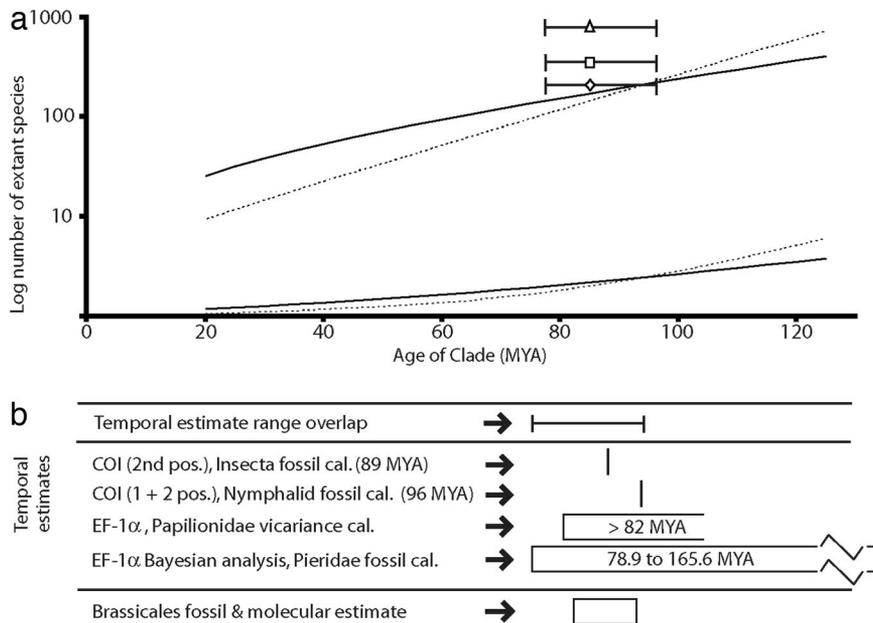
© 2007 by The National Academy of Sciences of the USA



**Fig. 1.** Chronogram of the Pieridae, showing host plant relationships and the NSP activity of representative species. Circles represent presence (filled) or absence (empty) of NSP activity in larvae from relevant genera. Vertical labels indicate higher taxa and informal groups. Branches are scaled relative to a divergence time of 85 Mya, shown on the x axis. Yellow branches refer to terminal taxa that are predominately Fabales feeding, whereas green branches are glucosinolate feeding, and blue are derived non-glucosinolate feeding (gray branches indicate unknown host plant associations). Pseudopontiinae and Dismorphiinae, sister clades to Coliadinae plus Pierinae, are not shown but can be seen in [SI Figs. 4 and 5](#). Higher taxon groupings are shaded for visualization. Butterfly images are of subfamily exemplar taxa. Red bar indicates the lower 95% confidence interval from an EF-1 $\alpha$  fossil calibrated, Bayesian relaxed molecular clock estimation of the Pierinae–Coliadinae divergence. Butterfly photos: M. Borsch (*A. crataegi*), S. Coombes (*A. cardamines*, *C. hyale*), M. Rowlings (*C. evagore*), H. Vogel (*P. rapae*).

mechanisms of the Lepidoptera: the cytochrome P450 monooxygenase gene superfamily in *Papilio* butterflies used against furanocoumarins (25) and the flavin-dependent monooxygenase system of the arctiid moth *Tyria jacobaeae* used against pyrrolizidine alkaloids (26).

With these issues in focus, we document a key innovation and its role in butterfly diversification by (i) assessing the evolutionary origin of a host plant detoxification mechanism involved in an ancient host plant shift, (ii) generating a robust estimate of when this detoxification mechanism evolved relative to the age



**Fig. 2.** Current and expected diversity, and divergence estimates for the Coliadinae and Pierinae subfamilies of Pieridae. (a) The x axis gives the age of clades (Mya) whereas the y axis gives the species numbers (Log scale) for each clade. The upper and lower 95% confidence limits are shown for the expected species numbers of a clade diversifying at a rate equal to that of the sister clade to Pierinae plus Coliadinae (i.e., Dismorphiinae plus Pseudopontiinae) with either no extinction (dashed lines) or a high extinction rate (90%, solid lines). Current diversity levels for Pierinae as a whole (triangle), the Pierinae glucosinolate feeders only (square), and Coliadinae (diamond) are plotted with an estimated origin 85 Mya. The horizontal bar within these groupings represents the overlap among the molecular divergence estimates shown below, using the fossil and molecular estimates of the age of the Brassicales as an upper bound. (b) COI and EF-1 $\alpha$  divergence estimates are scaled to the x axis shown in a. Single lines represent the COI point estimates. Rectangular boxes represent the range of EF-1 $\alpha$  vicariance estimates and the lower 95% confidence limit of Bayesian relaxed molecular clock estimates. Temporal estimates from Brassicales fossils are shown at bottom.

of the colonized host plant family, and (iii) testing for an effect of the host shift on butterfly diversification rate.

### Results

We begin by asking a mechanistic question, focusing on whether all Pierid glucosinolate feeders use NSP, not NSP, or the other identified glucosinolate detoxification mechanism of the Lepidoptera, GSS. Midgut assays of larvae from 13 Pieridae species from North America, Europe, Africa, and Australia, representing each of the major subfamilies or groups, reveals complete concordance between NSP activity and glucosinolate feeding (Fig. 1; and see supporting information (SI) Fig. 3 and SI Table 1). Consistent with initial studies of these mechanisms on *Pieris rapae*, none of the butterfly species assayed displays any detectable GSS activity, and *P. xylostella* and other moths have no NSP activity. Within the Pieridae, glucosinolate feeding exactly matches the presence of NSP activity in larvae, suggesting a single evolutionary origin with a subsequent loss after secondary host plant shifts to Santalales (Fig. 1, SI Fig. 3, and SI Table 1). These results support NSP as a key innovation within the Pieridae enabling glucosinolate detoxification (27).

Fossil and molecular data agree that the Brassicales appeared by 90 to 85 Mya, which is much earlier than the parallel evolution of glucosinolates in the Euphorbiaceae (58 Mya) (28). Pierinae fossils with modern relatives in the Brassicales feeding clade appear  $\approx 34$  Mya,  $\approx 50$  Myr after the first fossil Brassicales (13, 28). Fossils inherently provide only minimum temporal estimates. Thus, the fossil record is compatible with either a long delay of insect adaptation after the Brassicales radiation, a contemporaneous radiation of Brassicales and their Pierid herbivores, or anything in between. Among the major insect orders, Lepidoptera have the poorest fossil record, probably as a result of their wing structure (8, 75). In stark comparison, the fossil record of beetles is very rich, and beetles provide some of the best evidence for coevolution (29), although this evidence has

recently been questioned by a recalibrated phylogeny of the leaf beetles (30). In this article, we directly address the timing of the appearance of the glucosinolate-feeding Pierinae, using several independent molecular datasets and various calibration methods to generate a robust estimate of when Pieridae evolved relative to their Brassicales host plants.

Our temporal analysis used two standard molecular phylogenetic genes, cytochrome oxidase I (COI) and elongation factor 1 $\alpha$  (EF-1 $\alpha$ ), with two and three different temporal estimations, respectively (Fig. 2). First, we compared COI sequence from *Colias eurytheme* (Coliadinae) and *Pieris napi* (Pierinae) and estimated the time of their divergence at 89.12 and 96.32 Mya. These two estimates are respectively derived (i) from an Insecta-wide average substitution rate of 0.022% per million years for COI 2nd position data calibrated on ancient, nonlepidopteran fossils (31), and (ii) from COI 1st and 2nd position data calibrated on butterfly fossils from within the Nymphalinae subclade of Nymphalidae (32). Second, we compared the EF-1 $\alpha$  sequence distance between our Coliadinae and Pierinae representatives to the mean EF-1 $\alpha$  sequence distance between temporally calibrated representatives in the Nymphalidae and Papilionidae butterfly families (32, 33). Nymphalidae estimates were calibrated as stated above, and the Papilionidae samples were calibrated by using biogeographic vicariance events inferred during the breakup of Gondwana. Molecular evolution rates of the nuclear EF-1 $\alpha$  gene are likely to be similar among these families, as Nymphalidae and Papilionidae are Pieridae's nearest family clades (34). Together these EF-1 $\alpha$  distance comparisons suggest that Coliadinae diverged from Pierinae  $\gg 62$  and  $>82$  Mya, respectively (Fig. 2).

Pieridae-specific temporal reconstruction used a Bayesian relaxed molecular clock method on EF-1 $\alpha$  data from across Pieridae. Divergence estimates used the four known pierid fossils as minimum node dates (35, 36) (SI Figs. 4 and 5). Temporal reconstruction can be affected by assumptions used in analyses,

such as tree topology, node calibration, clock-like evolution, and other prior estimates (37). In our analysis, we explored the potential effects of alternative topologies, fossil node placements, clock assumptions, and other prior estimates on temporal inferences. These alternatives had little effects on results except to increase age estimates (SI Table 2). The most temporally conservative model predicts the appearance of Pieridae with a 95% confidence interval of 166 to 79 Mya, which agrees with related clock-based estimates (13). In sum, our five temporal estimates are consistent and provide a robust assessment. The glucosinolate-feeding Pierinae diverged from Coliadae in close evolutionary time to the appearance of the Brassicales, most likely within 10 million years (Figs. 1 and 2).

Ehrlich and Raven hypothesized, but were unable to test, that the “main diversification of [Pieridae] occurred after it became associated with the [Brassicales]” (8). Others have posited a relationship between host plant use and species abundance, but this relationship has not been tested (11, 12). We calculate and compare the standing diversity of the sister clades Coliadae and Pierinae (38), which are *a priori* expected to have similar levels of diversification because they are sister clades (39). There are both more genera and species of Pierinae than Coliadae (57 vs. 18 genera, 830 vs. 220 species). To assess diversification rate and extinction effects giving rise to these differences in extant taxa, we use the diversification rate of Dismorphiinae plus Pseudopontiinae (sister clade to Coliadae plus Pierinae) as the background rate for generating expected species diversity confidence intervals through time. Coliadae species diversity is at or below the expected numbers of species. However, there are significantly more species of Pierinae than expected, even if non-glucosinolate feeders are excluded, suggesting that the increased diversification of the Pierinae resulted from the host plant shift onto Brassicales (Fig. 2).

## Discussion

The evolution of the NSP glucosinolate detoxification gene was a novel, key innovation facilitating the host plant shift of ancestral Pierinae from Fabales to Brassicales feeding with macroevolutionary consequences. Our analysis of major clade representatives across Pieridae demonstrates that only glucosinolate-feeding Pierinae show NSP activity, indicating a single mechanistic basis of glucosinolate feeding originating within the Pierinae. Multiple independent temporal reconstructions provide a robust estimate placing the Pierinae host shift soon after the appearance of the Brassicales plant order. This host shift also resulted in significantly increased speciation compared with related clades. Together, these results provide strong support for the central tenet in Ehrlich and Raven’s coevolution theory, that key biochemical innovations foster increased speciation (8, 40).

Ancestral Pieridae butterflies did not provide the initial herbivory pressure for Brassicales chemical defense evolution, nor are they solely exerting sufficient pressure for the continued evolution of the glucosinolate-myrosinase system of the Brassicales, as lepidopteran host plant coevolution is expected to be diffuse (8, 40, 41). Yet, comparative functional genomic analysis of the glucosinolate-myrosinase system in the family Brassicaceae reveals ongoing evolution in response to herbivory pressure (42–44). Multiple enzymes, with allelic variation, are involved in the biosynthesis of diverse glucosinolate structures and relative concentrations (45–47). This resulting intra- and interspecific glucosinolate profile variation results in herbivory level variation (48–51). In addition, signatures of positive selection at the molecular level have been identified at these genes, which are undergoing gene duplication and neofunctionalization, documenting ongoing glucosinolate metabolic evolution (52). Thus, the initial innovation that likely facilitated Brassicales escape and radiation from herbivore pressure is continuing to evolve in derived lineages with herbivory consequences.

Studies of plant–insect coevolution have evolved significantly since the days of Ehrlich and Raven (e.g., refs. 53–64). Butterfly support for the parallel coevolutionary process, i.e., cospeciation, has proven difficult to find, as cospeciation-driven diversification seems to be rare among insects and butterflies in particular, with Ehrlich and Raven’s escape and radiate coevolutionary mechanism likely more common (8, 10, 40, 53, 65, 66). The only strong case for butterfly coevolution is found in the genus *Papilio* (Papilionidae), where  $\approx 75\%$  of this genus feeds on plants containing furanocoumarin-based chemical defenses, synthesized by plants using members of the cytochrome P450 monooxygenases gene superfamily (11, 25, 67). *Papilio* butterfly species diversity increases with host plant furanocoumarin diversity (11). Interestingly, furanocoumarins both induce and are detoxified by *Papilio* species using other members of the cytochrome P450 superfamily (68). Cytochrome P450 participation on both sides of the plant–insect interaction highlights the diverse functional roles members of this gene superfamily play in environmental response (68). In fact, the generalist Noctuid moth *Helicoverpa*, which is  $>100$  Myr divergent from *Papilio* sp., independently uses other P450s to detoxify furanocoumarins, highlighting the repeatability of this gene superfamily’s role in detoxification. Cytochrome P450s are also part of the glucosinolate biosynthetic pathway in Brassicales plants. Interestingly, in Lepidoptera specialized on glucosinolates, only the novel molecular mechanisms discussed in this article (GSS and NSP) have been identified. Given the apparently rapid host shift of the Pierinae onto Brassicales, we are investigating the evolutionary history of the NSP gene to assess its ancestral, pre-glucosinolate detoxification function.

Over 40 years ago, the authors who introduced us to coevolution using butterflies claimed that their “predictions cannot be tested” (8). The work presented here documents a key coevolutionary innovation in the Pieridae completely independent of other known lepidopteran detoxification mechanisms. Moreover, this work indicates a much more dynamic interaction between butterfly herbivores and their host plants than currently envisioned. We present our work here in the hope that evolutionary and ecological functional genomics will encourage increased mechanistic study of plant–insect interactions.

## Materials and Methods

**NSP and GSS Assay.** Last instar larvae actively feeding on preferred host plants (SI Table 1) were used to test for NSP activity using methods described in ref. 20. Briefly, benzylglucosinolate and myrosinase were incubated with larval midgut extracts. Dichloromethane extracts of these assay mixtures were then analyzed by GC-MS with internal standard, with total ion current traces identifying NSP activity (see SI Fig. 3 for representative traces and more details). Methods for detecting GSS activity in samples were as reported in ref. 24, although a much greater amount of protein (up to 100  $\mu$ g) was used in the assays in this study to detect potentially low enzyme activity levels.

**Phylogenetic Reconstruction.** Bayesian phylogenetic reconstruction, using a GTR + I + G nucleotide substitution model for a total of 1 million generations as implemented in Mr. Bayes v.3.1.2 (69, 70), was performed by using two different outgroups and different sets of the EF-1 $\alpha$  dataset of Braby *et al.* (13). Resulting tree 1 used *Papilio machaon* (Papilionidae) as an outgroup, without representatives *Pseudopontia* or Dismorphiinae (SI Fig. 4). Tree 2 used all available EF-1 $\alpha$  Pieridae sequences with *Pseudopontia paradoxa* as an outgroup (SI Fig. 5). Trees 1 and 2 place Dismorphiinae sister to Coliadae plus Pierinae, concordant with previous studies (13, 71). Tree topologies differed in their terminal branching patterns, whereas subfamily relationships were stable (see SI Materials and Methods for analysis details).

**Ancestral Node Dating. COI.** Recent analysis of Nymphalidae COI data, using fossil data for temporal calibration, estimated origins of the subgroup Nymphalinae at or older than 65 Mya (32). Using representative species *Melitaea* and *Vanessa*, from distant Nymphalinae clades, for 1st and 2nd position rate change per million year provides an estimate of 0.0004372 (26.8 changes per 943 bp;  $26.8/943 = 0.02841$ ;  $0.02841/65 = 0.0004372$  substitutions per million

years). There are 43 1st plus 2nd position changes between *C. eurytheme* and *P. napi*, giving a temporal estimate of the Coliadinae and Pierinae divergence at 96.32 Mya ( $43/1021 = 0.04211$ ;  $0.04211/0.0004372 = 96.32$ ).

**EF1 $\alpha$ .** The estimated EF-1 $\alpha$  mean genetic distance and standard deviation between Coliadinae ( $n = 15$ ) and Pierinae ( $n = 60$ ) taxa is 12.928% (SE = 0.0095), whereas that between *Papilio* species ( $n = 55$ ) and *Pachioptia nep-tunus* (Papilionidae) is nearly 3% lower at 10.029% (SE = 0.1115). Genetic distance was calculated by using a Tamura 3-Parameter distance model as implemented in Mega 3.1, with standard error calculated with 500 bootstrap replicates (72). Rate variation among sites was approximated by using a gamma shape parameter = 1.4 [as per the estimated value for the dataset (33)]. The two Papilionidae groups are calculated to have diverged 89.1 to 82.5 Mya, suggesting a comparable or even older divergence between Coliadinae and Pierinae. Similar EF-1 $\alpha$  calibrated estimates are available from Nymphalidae. By using representatives from *Vanessa* and *Melitaea* for a 65 million year split, there is a mean distance of 7.64% (SE = 0.0078). This distance is 60% of the divergence found between Coliadinae and Pierinae representatives, suggesting a Coliadinae Pierinae divergence much greater than 65 Mya.

**EF1 $\alpha$  using a relaxed clock method.** Temporal estimation using a Bayesian relaxed molecular clock method was implemented by using MULTIDIVTIME (35, 73) on the two Bayesian analysis EF-1 $\alpha$ -generated trees (SI Figs. 4 and 5). Prior estimates (SI Table 2) explored model response across many different prior values following a similar approach by Wiegmann et al. (74). To perform the most conservative analysis possible, all analyses used a prior maximum tree age set to the oldest known fossil Pierinae at 34 Mya. For all combinations of prior values, the Markov Chain was then sampled every 100 cycles for a total of 10,000 samples. Posterior distributions were approximated based on these 10,000 samples. Fossil node placement initially followed Braby et al. (13), but here fossils are assigned as a minimum date estimate and the effect of more conservative node placement was explored by using the next most ancestral node (SI Figs. 4 and 5 and SI Table 2). For each run, across two trees, the prior conditions, resulting estimates, and confidence limits are listed in SI Table 2.

**Absolute Diversification Rate Calculations.** There are both more genera and species of Pierinae than Coliadinae (57 vs. 18 genera; 830 vs. 220 species). These numbers are significantly different from equal number expectations ( $\chi^2 = 10.8$ ,  $P = 0.001$ ;  $\chi^2 = 193.5$ ,  $P < 0.001$ ). However, this simple calculation does not assess diversification rate and extinction effects giving rise to these differences in extant taxa. To address these effects, we perform a statistical analysis of a stochastic birth-and-death processes, employing a method-of-moments estimator following Magallon and Sanderson (38), which takes the age of the clades in question, together with their extant species diversities, and allows for them to be compared with their expected levels of diversification given different extinction rate assumptions (see SI Materials and Methods for more details). Sister clades Coliadinae and Pierinae are by definition of equal age, and, for an unbiased, phylogenetically relevant estimator of their expected species diversity, we chose the diversification rate of their sister clade, Dismorphiinae plus Pseudopontiinae. By using this rate, 95% confidence limits were calculated for expected numbers of species across different clade ages with either no extinction or a high (90%) extinction rate (see SI Materials and Methods for more details; SI Table 3). Groups with a greater than expected number of species are considered to be “excessively species rich” (38). The 95% confidence interval for high extinction is generally greater than under no extinction due to having a higher variance; a high number of speciation events are needed to offset the high number of extinctions to maintain the same net diversification rate as when there is no extinction.

**ACKNOWLEDGMENTS.** We thank K. Bargum, M. Clauss, J. Marden, S. Nylin, R. Schilder, N. Wahlberg, and W. Watt for comments on various stages of this manuscript; R. Oyama and J. Thorne for analysis assistance; and A. Bonkewitz and D. Heckel for specimen collection help. M. Borsch, S. Coombes, and M. Rowlings kindly permitted our use of their butterfly photographs (*A. crataegi*, *A. cardamines*, and *C. hyale* and *C. evagore*, respectively). This work was funded by the Max Planck Society and National Science Foundation Grant IBN-0412651 (to C.W.W.).

- Labandeira CC, Dilcher DL, Davis DR, Wagner DL (1994) *Proc Natl Acad Sci USA* 91:12278–12282.
- Vane-Wright D (2004) *Nature* 428:477–478.
- Magallon SA, Sanderson MJ (2005) *Evolution (Lawrence, Kans)* 59:1653–1670.
- de Jong R (2003) *Invertebr Syst* 17:143–156.
- Braby MF, Trueman JWH, Eastwood R (2005) *Invertebr Syst* 19:113–143.
- Miller JY, Miller LD (2001) in *Biogeography of the West Indies: Patterns and Perspectives*, eds Woods CA, Sergile FE (CRC Press, Boca Raton, FL), pp 127–150.
- Labandeira CC, Sepkoski JJ (1993) *Science* 261:310–315.
- Ehrlich PR, Raven PH (1964) *Evolution (Lawrence, Kans)* 18:586–608.
- Wahlberg N, Brower AVZ, Nylin S (2005) *Biol J Linn Soc* 86:227–251.
- Janz N, Nylin S (1998) *Evolution (Lawrence, Kans)* 52:486–502.
- Berenbaum M (1983) *Evolution (Lawrence, Kans)* 37:163–179.
- Braby MF, Trueman JWH (2006) *J Evol Biol* 19:1677–1690.
- Braby MF, Vila R, Pierce NE (2006) *Zool J Linn Soc* 147:238–275.
- Hall JC, Iltis HH, Sytsma KJ (2004) *Syst Bot* 29:654–669.
- Chew F (1988) in *Biologically Active Natural Products*, ed Culter HG (Am Chem Soc, Washington, DC), pp 155–181.
- Mitchell-Olds T, Clauss MJ (2002) *Curr Opin Plant Biol* 5:74–79.
- Wittstock U, Halkier BA (2002) *Trends Plant Sci* 7:263–270.
- Koroleva OA, Davies A, Deeken R, Thorpe MR, Tomos AD, Hedrich R (2000) *Plant Physiol* 124:599–608.
- Andreasen E, Jorgensen LB, Høglund AS, Rask L, Meijer J (2001) *Plant Physiol* 127:1750–1763.
- Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T, Gershenson J, Vogel H (2004) *Proc Natl Acad Sci USA* 101:4859–4864.
- Li Q, Eigenbrode SD, Stringham GR, Thiagarajah MR (2000) *J Chem Ecol* 26:2401–2419.
- Lichtenstein EP, Morgan DG, Müller CH (1964) *J Agric Food Chem* 12:158–161.
- Wittstock U, Kliebenstein DJ, Lambrix V, Reichelt M, Gershenson J (2003) in *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology*, Recent Advances in Phytochemistry, eds Romeo JT (Elsevier, Amsterdam), Vol 37, pp 101–125.
- Ratzka A, Vogel H, Kliebenstein DJ, Mitchell-Olds T, Kroymann J (2002) *Proc Natl Acad Sci USA* 99:11223–11228.
- Li WM, Schuler MA, Berenbaum MR (2003) *Proc Natl Acad Sci USA* 100:14593–14598.
- Naumann C, Hartmann T, Ober D (2002) *Proc Natl Acad Sci USA* 99:6085–6090.
- Berenbaum MR, Favret C, Schuler MA (1996) *Am Nat* 148:5139–5155.
- Wikstrom N, Savolainen V, Chase MW (2001) *Proc R Soc London Ser B* 268:2211–2220.
- Farrell BD (1998) *Science* 281:555–559.
- Gómez-Zurita J, Hunt T, Koplíku F, Vogler AP (2007) *PLoS ONE* 1:e360.
- Gaunt MW, Miles MA (2002) *Mol Biol Evol* 19:748–761.
- Wahlberg N (2006) *Syst Biol* 55:703–714.
- Zakharov EV, Caterino MS, Sperling FAH (2004) *Syst Biol* 53:193–215.
- Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, Nylin S, Pierce NE, Sperling FAH, Vila R, Warren AD, Zakharov E (2005) *Proc R Soc London Ser B* 272:1577–1586.
- Kishino H, Thorne JL, Bruno WJ (2001) *Mol Biol Evol* 18:352–361.
- Thorne JL, Kishino H, Painter IS (1998) *Mol Biol Evol* 15:1647–1657.
- Ho SYW, Phillips MJ, Drummond AJ, Cooper A (2005) *Mol Biol Evol* 22:1355–1363.
- Magallon S, Sanderson MJ (2001) *Evolution (Lawrence, Kans)* 55:1762–1780.
- Mitter C, Farrell B, Wiegmann BM (1988) *Am Nat* 132:107–128.
- Futuyma DJ (2000) *Plant Species Biol* 15:1–9.
- Hougen-Eitzman D, Rausher M (1994) *Am Nat* 143:677–697.
- Kliebenstein DJ, Kroymann J, Mitchell-Olds T (2005) *Curr Opin Plant Biol* 8:264–271.
- Raybould AF, Moyes CL (2001) *Heredity* 87:383–391.
- Renwick JAA (2002) *Entomol Exp Appl* 104:35–42.
- de Quiros HC, Magrath R, McCallum D, Kroymann J, Scnabelrauch D, Mitchell-Olds T, Mithen R (2000) *Theor Appl Genet* 101:429–437.
- Textor S, Bartram S, Kroymann J, Falk KL, Hick A, Pickett JA, Gershenson J (2004) *Planta* 218:1026–1035.
- Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenson J, Mitchell-Olds T (2001) *Plant Physiol* 127:1077–1088.
- Heidel AJ, Clauss MJ, Kroymann J, Savolainen O, Mitchell-Olds T (2006) *Genetics* 173:1629–1636.
- Windsor AJ, Reichelt M, Figuth A, Svatos A, Kroymann J, Kliebenstein DJ, Gershenson J, Mitchell-Olds T (2005) *Phytochemistry (Amsterdam)* 66:1321–1333.
- Kroymann J, Donnerhacke S, Schnabelrauch D, Mitchell-Olds T (2003) *Proc Natl Acad Sci USA* 100:14587–14592.
- Mauricio R (1998) *Am Nat* 151:20–28.
- Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenson J, Kroymann J (2006) *Proc Natl Acad Sci USA* 103:9118–9123.
- Pellmyr O (2003) *Ann Mo Bot Gard* 90:35–55.
- Forde SE, Thompson JN, Bohannan BJM (2004) *Nature* 431:841–844.
- Cornell HV, Hawkins BA (2003) *Am Nat* 161:507–522.
- Thompson JN, Fernandez CC (2006) *Ecology* 87:103–112.
- Thompson JN (2005) *Curr Biol* 15:R992–R994.
- Thompson JN (2005) *Geographic Mosaic of Coevolution* (Univ of Chicago Press, Chicago).
- Scriber JM (2002) *Entomol Exp Appl* 104:217–235.
- Becerra JX (1997) *Science* 276:253–256.
- Becerra JX (2003) *Proc Natl Acad Sci USA* 100:12804–12807.
- Becerra JX, Venable DL (1999) *Proc Natl Acad Sci USA* 96:12626–12631.
- Thompson JN, Cunningham BM (2002) *Nature* 417:735–738.
- Nuismer SL, Thompson JN (2006) *Evolution (Lawrence, Kans)* 60:2207–2217.
- Janz N, Nylin S, Wahlberg N (2006) *BMC Evol Biol* 6:4.
- Edwards SV, Beerli P (2000) *Evolution (Lawrence, Kans)* 54:1839–1854.
- Berenbaum MR (2001) *Ann Mo Bot Gard* 88:45–59.
- Berenbaum MR (2002) *J Chem Ecol* 28:873–896.
- Huelsenbeck JP, Ronquist F (2001) *Bioinformatics* 17:754–755.
- Ronquist F, Huelsenbeck JP (2003) *Bioinformatics* 19:1572–1574.
- Pollock DD, Watt WB, Rashbrook VK, Iyengar EV (1998) *Ann Entomol Soc Am* 91:524–531.
- Kumar S, Tamura K, Nei M (2004) *Brief Bioinform* 5:150–163.
- Thorne JL, Kishino H (2002) *Syst Biol* 51:689–702.
- Wiegmann BM, Yeates DK, Thorne JL, Kishino H (2003) *Syst Biol* 52:745–756.
- Kristensen NP, Scoble MJ, Karsholt O (2007) *Zootaxa*, in press.