MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 340–349 doi: 10.1111/j.1365-294X.2011.05379.x

The evolution of cardenolide-resistant forms of Na^+,K^+-ATPase in Danainae butterflies

MATTHEW L. AARDEMA,* YING ZHEN*† and PETER ANDOLFATTO*†

*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA, †Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

Abstract

Cardenolides are a class of plant secondary compounds that inhibit the proper functioning of the Na^+, K^+-ATPase enzyme in susceptible animals. Nonetheless, many insect species are able to sequester cardenolides for their own defence. These include butterflies in the subfamily Danainae (Family: Nymphalidae) such as the monarch (*Danaus plexippus*). Previous studies demonstrated that monarchs harbour an asparagine (N) to histidine (H) substitution (N122H) in the α subunit of Na^+, K^+-ATPase (ATPα) that reduces this enzyme's sensitivity to cardenolides. More recently, it has been suggested that at ATPα position 111, monarchs may also harbour a leucine (L)/glutamine (Q) polymorphism. This later amino acid could also contribute to cardenolide insensitivity. However, here we find that incorrect annotation of the initially reported DNA sequence for ATPα has led to several erroneous conclusions. Using a population genetic and phylogenetic analysis of monarchs and their close relatives, we show that an ancient Q111L substitution occurred prior to the radiation of all Danainae, followed by a second substitution at the same site to valine (V), which arose before the diversification of the *Danaus* genus. In contrast, N122H appears to be a recent substitution specific to monarchs. Surprisingly, examination of a broader insect phylogeny reveals that the same progression of amino acid substitutions (Q111L → L111V + N122H) has also occurred in *Chrysochus* beetles (Family: Chrysomelidae, Subfamily: Eumolpinae) that feed on cardenolide-containing host plants. The parallel pattern of amino acid substitution in these two distantly related lineages is consistent with an adaptive role for these substitutions in reducing cardenolide sensitivity and suggests that their temporal order may be limited by epistatic interactions.

Keywords: cardiac glycosides, Danainae, *Danaus plexippus*, milkweed, molecular adaptation, monarch butterfly, parallel evolution

Received 9 August 2011; revision received 16 October 2011; accepted 21 October 2011

Introduction

Many plants are defended from consumption by secondary metabolites that are toxic to invertebrate herbivores. Nonetheless, numerous insect species have evolved physiological adaptations that allow them to overcome these defences and feed on chemically protected plant tissues without apparent harm. It has been proposed that the evolution of such abilities has lead to the radiation of multiple insect taxa (e.g. Ehrlich & Raven 1964; Futyuma & Keese 1992; Berenbaum *et al.* 1996; Kergoat *et al.* 2005; Wheat *et al.* 2007). Well-studied examples of such radiations include insects that feed on members of the plant family Apocynaceae (Farrell & Mitter 1994, 1998; Agrawal *et al.* 2009). Many of the plants in this family produce toxic cardiac glycosides (CGs) or cardenolides that help to deter herbivores. In a susceptible animal, cardenolides bind to the α subunit of Na^+, K^+-ATPase (ATPα), inhibiting its function and ultimately affecting many important physiological processes including muscle contraction, neural function and ion transport (Schatzmann 1964; Holzinger & Wink 1996). Nevertheless, several groups of specialist

Correspondence: Matthew L. Aardema, Fax: +1 609 258 1712; E-mail: maardema@princeton.edu

insects feed on cardenolide-producing plants and some additionally sequester plant-derived CG compounds to aid in their own defence (Duffey & Scudder 1972; Brower & Glazier 1975; Cohen & Brower 1983; Ackery & Vane-Wright 1985; Dobler *et al.* 1998; Nishida 2002). An example of this are butterflies in the subfamily Danainae (family: Nymphalidae) that as larvae feed on cardenolide-rich plants such as those in the genus *Asclepias* (family: Apocynaceae). Some of the butterflies in this clade, specifically those in the genus *Danaus* such as the monarch (*Danaus plexippus*), are also able to sequester CG compounds in their wings and integument (Ackery & Vane-Wright 1985; Frick & Wink 1995; Mebs *et al.* 2005) where they deter consumption by predators (Brower 1958).

ATPα has been examined in many organisms including the monarch to determine whether specific amino acid substitutions may contribute to cardenolide insensitivity and sequestration ability (e.g. Canfield *et al.* 1992; Holzinger *et al.* 1992; Jaisser *et al.* 1992; Holzinger & Wink 1996; Mebs *et al.* 2000; Labeyrie & Dobler 2004; Zhu *et al.* 2008). Site-directed mutagenesis (reviewed in Croyle *et al.* 1997) and protein structure studies (Ogawa *et al.* 2009; Yatime *et al.* 2011) have identified numerous amino acid residues throughout ATPα that interact with cardenolides such as ouabain (which is similar to those naturally occurring in many Apocynaceae plants; Frick & Wink 1995). Variation in a 12 amino acid extracellular domain located between the H1 and H2 transmembrane regions of ATPα drastically influences the enzyme's sensitivity to ouabain (e.g. Price & Lingrel 1988; Price *et al.* 1989, 1990; Canfield *et al.* 1990; Jewell & Lingrel 1992). Accordingly, changes here are suspected of contributing to cardenolide insensitivity in monarchs, even though other Danainaes are believed to lack substitutions in this region (Holzinger *et al.* 1992; Frick & Wink 1995; Holzinger & Wink 1996; Mebs *et al.* 2000, 2005).

Specifically, monarchs are reported to have an asparagine (N) to histidine (H) substitution (N122H) (Holzinger *et al.* 1992; Holzinger & Wink 1996), which reduces cardenolide sensitivity experimentally (Holzinger & Wink 1996). Another study suggests that monarchs may harbour a glutamine (Q)/valine (V) polymorphism at a second position, 111 (Zhu *et al.* 2008). Such a change could also contribute to a reduction in sensitivity to CG compounds (Price & Lingrel 1988; Price *et al.* 1990). However, other Danainaes are reported to lack substitutions in the H1–H2 extracellular domain (Holzinger & Wink 1996; Mebs *et al.* 2000), raising the question of how these butterflies manage to avoid the possible detrimental effects inherent to consuming cardenolide-containing plant tissue.

A careful examination of published accounts for monarchs and other Danainae butterflies suggested to us that the information concerning the H1–H2 extracellular domain might be inaccurate and/or incomplete. Specifically, we found that the annotated intron/exon splice sites in the published sequences do not conform to well-established intron splice junction sequences. Here we re-examine previously published information and present new population genetic and phylogenetic data concerning the nature and temporal order of amino acid substitutions in the H1–H2 extracellular domain of Danainae ATPα. Our results have implications for the genetic basis of cardenolide insensitivity in resistant insects and, more generally, models of adaptive protein evolution.

Methods

Re-examination of previously published sequences

We obtained published DNA sequence data for the H1–H2 extracellular domain of ATPα from the following organisms: *Drosophila melanogaster* (GenBank accession no.: BT125846), *Bombyx mori* (Silkbase ID: BGIBMGA005058-TA), *D. plexippus* (GenBank accession no.: S51586) and the queen butterfly, *D. gilippus* (Holzinger & Wink 1996). Using these sequences, we examined the published annotations of the H1–H2 extracellular domain for monarchs and queens (Holzinger *et al.* 1992; Holzinger & Wink 1996), noting the presence of an incorrectly annotated intron in both sequences. Where possible, we re-annotated these sequences so that the intron began and ended with consensus splice sites 5′ G|GTRAGT and 3′ CAG|G (Lewin 2004).

cDNA sequencing of the ouabain-binding domain

To confirm the correct annotation of the intron within the H1–H2 extracellular domain of ATPα, we PCR-amplified this region from cDNA isolated from one monarch and one queen individual. The monarch was collected from Charlevoix County, Michigan, and the queen from Jefferson County, Florida, both in 2010. Thoracic tissue was dissected and stored in RNAlater (Qiagen) at −80 °C. Total RNA was extracted using the Qiagen RNeasy mini kit (Qiagen) and was treated with RNase-free DNase (Qiagen). Following first-strand cDNA synthesis using oligo $(dT)_{18}$ primer (M-MLV Reverse Transcriptase, Promega), we PCR-amplified the H1–H2 extracellular domain using primers designed from the published monarch EST sequence (GenBank accession no.: EY260360; forward: 5′-GCGAAAGA GAACCTTGAACG-3′; reverse: 5′-GTCCACGAGACT GAGGTTCC-3′). The reaction conditions were an initial incubation at 95 °C for 4 min, followed by 35 cycles of

95 °C for 30 s, 55 °C for 1 min and 72 °C for 1 min, and then incubation at 72 °C for 7 min. Sanger-based sequencing was carried out using the same primers and run on an Applied Biosystems Automated 3730 DNA Analyzer at the Life Sciences Core Laboratories Center (Cornell University, Ithaca, NY, USA).

Genomic sequencing of ouabain-binding domain

To look for potentially important amino acid polymorphisms in the H1–H2 extracellular domain of ATPα, we sequenced genomic DNA from a panel of 69 monarchs and seven queen butterflies (Table S1, Supporting information). Genomic DNA was prepared using a standard proteinase-K, phenol–chloroform protocol. We PCR-amplified and sequenced the region spanning the H1–H2 extracellular domain of ATPα using the same primers and PCR conditions used for cDNA (above). These primers were designed to span the predicted intron in the extracellular region. The resulting PCR products were purified and Sanger-sequenced as described earlier. Sequence alignment and annotation were performed using SEQUENCHER 4.7 (Gene Codes Corp.).

We additionally sequenced this region of the protein in five other members of the subfamily Danainae that may feed on cardenolide-containing host plants (Ackery 1988), as well as from a non-cardenolide utilizing species, *Limenitis archippus*, in the Nymphalidae family (Table S1, Supporting information). As the DNA prepared from the additional Danainaes was from dried samples, we utilized a primer pair that amplified a shorter region of the gene, but still spanned the extracellular region, again designed from the published monarch EST sequence (forward: 5′-TCTCTTCGGTGGTTTCGCGT-3′; reverse: 5′-GTCACGATAACGACAGCCGC-3′). Reaction conditions were the same as those described earlier.

Polymorphism and divergence

Recent mutations in the H1–H2 extracellular domain of ATPα that contribute to increased cardenolide resistance or sequestration ability could have resulted in a selective sweep in this region, reducing nucleotide diversity. We tested this possibility by examining levels of diversity in monarchs and queens, as well as divergence between them, using the intronic region that lies within the H1–H2 extracellular domain between amino acids 111 and 112 of ATPα. After omitting the GT/AG splice sites, this intron is 85 nucleotides long in monarchs and 77/76 nucleotides long in queens. Insertion/deletion sites (indels) were excluded from our analysis of divergence between these species.

We calculated derived allele frequencies as the average frequency of derived mutations across all sites. The derived state was determined by parsimony using *Tirumala petiverana* as an outgroup (Table S2, Supporting information). All sites could be assigned unambiguously. Polymorphism within each species was calculated using two methods: the average pairwise diversity per site (π) and Watterson's (1975) diversity estimator (θ_W). We calculated θ_W using the number of segregating mutations, as one site in monarchs had three different segregating alleles. Tajima's *D* (Tajima 1989) was calculated using a custom script. We estimated the *P*-value of the observed Tajima's *D* under the standard neutral model using the coalescent program MS (Hudson 2002). Lineage-specific divergence (D_{XY}) was calculated as the average pairwise number of nucleotide substitutions per site along a lineage (Nei 1987) with Jukes–Cantor correction (Jukes & Cantor 1969).

Broader phylogenetic examination

To better understand the evolution of amino acid substitutions in the H1–H2 extracellular domain of ATPα, we aligned our amino acid sequences (plus two additional Lepidopteran species) to those of members of the other three major holometabolous (complete metamorphosis) insect orders (Diptera, Coleoptera and Hymenoptera) and a hemimetabolous (incomplete metamorphosis) insect outgroup (order: Orthoptera; Fig. 1; Table S3, Supporting information). These sequences included the 12 amino acids of the H1–H2 extracellular domain, plus 14 upstream residues and 10 downstream residues. Of particular interest were beetles in the genus *Chrysochus* (family: Chrysomelidae, subfamily: Eumolpinae) some of which are reported to feed on cardenolide-containing host plants and exude cardenolides in defensive secretions (Dobler *et al.* 1998; Dobler & Farrell 1999; Dobler 2004). We constructed a phylogenetic tree showing species relationships to one another, using data from several published sources (Fig. 1; Table S3, Supporting information; Flook & Rowell 1998; Regier *et al.* 1998; Labeyrie & Dobler 2004; Cameron *et al.* 2007; Hunt *et al.* 2007; Reidenbach *et al.* 2009; Wahlberg *et al.* 2009; Wiegmann *et al.* 2009; Brower *et al.* 2010). Amino acid alignment was carried out using SEQUENCHER 4.7 (Gene Codes Corp.).

Results

Re-examination of previously published sequences

Several published sources provide sequence data and a proposed annotation for an intron near the H1–H2 extracellular domain of ATPα in Danainae butterflies (Holzinger *et al.* 1992; Holzinger & Wink 1996; Mebs *et al.* 2000;

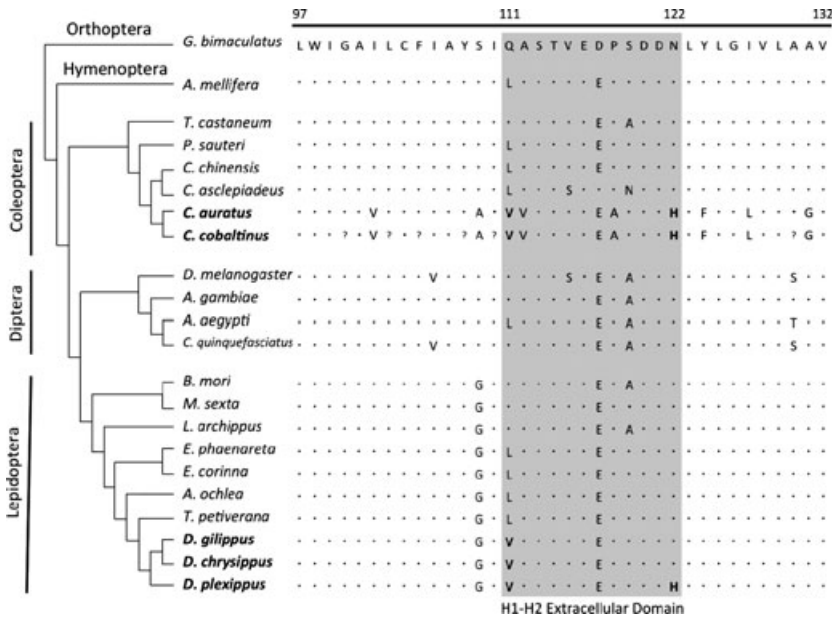

Fig. 1 Phylogeny and amino acid sequences of the H1–H2 extracellular domain and adjacent regions in the Lepidopteran species sequenced for this study, plus additional insects from the four major holometabolous insect orders (Lepidoptera, Diptera, Coleoptera & Hymenoptera), as well as one representative outgroup species (order: Orthoptera). Species that are reported to sequester cardenolides are bolded, as are important amino acid changes discussed in the text. The presence of a valine at position 111 in the three *Danaus* butterflies as well as *Chrysochus auratus* and *Chrysochus cobaltinus* may indicate parallel evolution of cardenolide sequestration ability under mutational constraints and epistatic interactions. In both cases, a mutation from glutamine to leucine may have facilitated an eventual replacement by valine at this position. The N122H mutation seen in monarchs and the two *Chrysochus* beetles likely contributes to cardenolide sequestration ability. See text for additional details.

Zhu *et al.* 2008). Holzinger *et al.* (1992) originally placed the intron between amino acids 109 and 110 in the α subunit (positions 43–131 in GenBank accession no. S51586), immediately before the H1–H2 extracellular domain. With this proposed placement of the intron, the first amino acid of this domain is inferred to be glutamine, which is identical to other Lepidoptera that do not feed on cardenolide-containing host plants (Holzinger *et al.* 1992). However, this reported intron sequence begins with CT and ends with TC, which does not conform to known splicing patterns. We thus re-annotated the intron (to positions 46–134 in GenBank accession no. S51586) so that it begins and ends with consensus splice sites 5′ G|GTRAGT and 3′ CAG|G, respectively (Lewin 2004). Assuming this DNA sequence is correct, the H1–H2 extracellular domain of the resulting protein for this monarch sequence starts with leucine at position 111 instead of glutamine.

In addition to the monarch, queen butterflies have also been examined for amino acid substitutions in the H1–H2 extracellular domain (Holzinger & Wink 1996). Following the annotation of the intron proposed by Holzinger *et al.* (1992), Holzinger & Wink (1996) concluded that queens also have a glutamine at position 111. As discussed earlier, the correct placement of the intron in monarchs is between positions 111 and 112. However, using the published sequence data for queens (Holzinger & Wink 1996), placement of the intron at this location leads to problems because the intron still does not begin with the characteristic splice junction nucleotides GT (although it does end with AG). If we insert the intron in a way that conforms to the consensus 5′ G|GTRAGT and 3′ CAG|G intron boundaries, the only options result in frameshift mutations. Based on our own sequencing of this region in queen butterflies, we conclude that a frameshift is unlikely and rather that the reported nucleotide sequence is incorrect.

Holzinger & Wink (1996) provide sequences and annotations for four additional Lepidopteran species, *Manduca sexta*, *Creatonotos transiens*, *Syntomeida epilais* and *Syntomis mogadorensis*. Despite the fact that the intron was incorrectly annotated in three of these species, correcting this error coincidentally places a glutamine at position 111, as originally proposed. For the fourth species, *S. epilais*, the sequence provided is not sufficient to determine the correct annotation.

cDNA sequencing of the ouabain-binding domain

Given the uncertainty regarding the previously published sequences, we sequenced cDNA from one monarch and one queen butterfly to confirm the correct

annotation and sequence of the H1–H2 extracellular region of ATPα. For the monarch, our results revealed valine at position 111, rather than leucine as predicted by re-annotation of the previously published sequence (S51586). We also found a histidine at position 122 in monarchs, as previously reported. The queen individual we sampled lacks the histidine at position 122 but shares the valine at position 111, suggesting that the former substitution is specific to monarchs, whereas the latter substitution predates divergence of the monarch and queen lineages.

Genomic sequencing of monarchs and other species

As previous studies reported glutamine at position 111 in monarchs (Holzinger *et al.* 1992; Holzinger & Wink 1996; Mebs *et al.* 2000), Zhu *et al.* (2008) interpreted their finding of valine at this position as a potential amino acid polymorphism. As we have shown, the original amino acid sequence was based on an incorrectly annotated intron. However, correct annotation places a leucine at position 111, which still suggests that a leucine/valine polymorphism may exist in monarchs. We tested this possibility by sequencing genomic DNA from the H1–H2 extracellular domain in 69 monarchs, collected from a variety of locations and years (Table S1, Supporting information). Corroborating reported sequences, we found an 89 base pair intron in all individuals and each of these had the consensus 5′ G|GTRAGT and 3′ CAG|G splice sites and valine at position 111. No evidence for an amino acid polymorphism at this position was found. We conclude that if a leucine/valine polymorphism does exist in monarchs, it likely occurs at very low frequencies. Considering the reliability of sequencing technology available to Holzinger *et al.* (1992), sequencing error should be considered as a possible alternative explanation.

We also included seven queen butterflies in our genomic DNA sequence survey. Previous studies suggested that queens have glutamine at position 111, and it was unclear how to correctly splice the reported intron sequence. Our genomic sequencing of queens revealed an 80-81 base pair intron in all individuals with consensus splice sites 5′ G|GTRAGT and 3′ CAG|G. The queen intron differs from that of monarchs by three fixed insertion-deletion (indel) differences: two, one base pair indels and one, nine base pair indel. In addition, queens are polymorphic for a one base pair indel in this intron. More importantly, we found that all sequenced queens possess valine at position 111, consistent with our cDNA sequence. This provides further evidence that the L111V substitution is likely fixed in both species and predates the divergence of the monarch and queen lineages.

To better assess larger evolutionary patterns of substitution in the H1–H2 extracellular domain of ATPα, we sequenced this region in a third *Danaus* species (*Danaus chrysippus*), four additional Danainaes (*Tirumala petiverana*, *Amauris ochlea*, *Euploea corinna* and *Euploea phaenareta*) and *Limenitis archippus*, which does not naturally encounter cardenolides (Ritland & Brower 1993). *Danaus chrysippus* has valine at position 111, similar to the monarchs and queens surveyed. The other Danainaes have leucine at position 111 and *L. archippus* has a glutamine, like other non-Danainae Lepidoptera (Fig. 1). These results imply that a L111V mutation likely arose before the diversification of the Danainae butterflies (~70 mya; Zhang *et al.* 2008) and was preceded by a more ancient substitution of Q111L that preceded the diversification of the *Danaus* genus (~5 Mya; Lushai *et al.* 2003). In contrast to these ancient substitutions, N122H is specific to monarchs, implying that it arose sometime after the speciation of queens and monarchs.

Polymorphism and divergence

Given the possibility of a recent selective sweep associated with the N122H substitution in monarchs, we examined polymorphism patterns in the short intron that interrupts the H1–H2 extracellular domain in both monarch and queens (Tables 1 and S2, Supporting information). Surprisingly, we found that monarchs actually have a higher level of nucleotide variation in this intron than queens (Table 1). For both species, Tajima's *D* was somewhat negative but not significantly different from expectations under the standard neutral model ($P > 0.05$ by neutral coalescent simulation). In addition, we see a number of intermediate frequency variants in monarchs with no obvious linkage disequilibrium among them. This suggests that the N122H substitution, which was likely to be adaptive (Holzinger & Wink 1996), was either not relatively recent or not selected strongly enough to create a detectable selective sweep signature in this genomic region.

Broader phylogenetic examination

The H1–H2 extracellular region, like the rest of ATPα, is highly conserved with few amino acid changes between holometabolous insects and a hemimetabolous outgroup which diverged approximately 355 Ma (Wiegmann *et al.* 2009). Layberie and Dobler (2004) documented that the N122H substitution occurs both in the monarch and in cardenolide-sequestering *Chrysochus* beetles (*Chrysochus auratus* and *Chrysochus cobaltinus*). However, based on our new data and analysis, we can see that there have actually been two additional parallel amino acid substitutions in the Danainae and Eumolpi-

Table 1 Polymorphism and divergence statistics for monarch and queen butterflies

Species	No. of individuals	Number of sites	S*	θ_W†	π‡	D§	D_{XY}¶	Tajima's D
Monarch	69	85	20	0.0428	0.0295	5	0.0774	−0.8643
Queen	7	77/76**	5	0.0207	0.0185	3	0.0503	−0.3647

*Total number of observed polymorphisms.
†Estimate of θ per site (Watterson 1975).
‡Average pairwise diversity per site.
§Number of fixed nucleotide substitutions.
¶Lineage-specific, Jukes–Cantor corrected estimates of nucleotide divergence per site.
**Queens are polymorphic for a 1 base pair indel (see text).

nae subfamilies. From the phylogeny (Fig. 1), it is clear that Q111L arose prior to the diversification of *Chrysochus* beetles and, like in Danainae, this was followed by a L111V substitution in the lineage leading to *C. auratus* and *C. cobaltinus*, both of which sequester cardenolides. These results imply that the same order of substitutions (Q111L → L111V + N122H) occurred independently in both of these cardenolide-sequestering taxa.

Discussion

Our re-examination of the originally reported monarch genomic sequence of ATPα suggests a leucine is present at position 111 instead of the previously reported glutamine (Holzinger *et al.* 1992). Glutamine is found at this position in Lepidoptera that do not feed on cardenolide-containing host plants such as *B. mori*, *M. sexta* and *L. archippus* (Fig. 1). A glutamine is likely the ancestral state in Lepidoptera and possibly all insects. A larger phylogenetic examination suggests that the L111V substitution preceded the diversification of the Danainaes and that this was preceded by an ancient Q111L substitution. We conclude that the leucine observed in the monarch sequence of Holzinger *et al.* (1992) may represent either a rare amino acid polymorphism or sequencing error.

Interestingly, in both insect lineages examined with members capable of sequestering cardenolides (subfamilies: Danainae & Eumolpinae), we observe the same temporal sequence of amino acid changes at this site (i.e. Q111L followed by L111V). Only single-nucleotide mutations are needed to make each of these two transitions in turn. It is not possible to go directly from glutamine to valine with only a single-nucleotide change. The transition from glutamine to leucine may be common and easily facilitated as two additional surveyed species, *Apis mellifera* (order: Hymenoptera) and *Aedes aegypti* (order: Diptera), also have a leucine at this position in the H1–H2 extracellular domain (Fig. 1). Although the functional significance of the Q111L and L111V substitutions is unclear, their appearance in both the Danainae and Eumolpinae subfamilies is suggestive of a functional role in reducing cardenolide sensitivity of ATPα. Several other types of amino acid substitutions at this position (notably Q111R, Q111H and Q111D) have been shown to at least partially reduce the sensitivity of this enzyme to ouabain (Price *et al.* 1990). Although no mutagenesis studies have looked specifically at the lone effects of either Q111L or L111V, the former substitution reduces ouabain sensitivity when combined with N122D (reviewed in Croyle *et al.* 1997).

A similar mutation, N122H, has been shown to reduce the sensitivity of ATPα to cardenolides and may facilitate cardenolide sequestration (Price *et al.* 1990; Holzinger & Wink 1996). This mutation is observed in both cardenolide-sequestering *Chrysochus* beetles and the monarch, but not in other *Danaus* spp., implying that it is recently derived in these butterflies. However, our population genetic analysis using the closely linked intron of the H1–H2 extracellular domain found no evidence of a recent selective sweep in monarchs. Signatures of selective sweeps are expected to be short-lived, on the order of N_e generations (Przeworski 2002). If the N122H substitution occurred during or relatively shortly after speciation, it is not surprising that evidence of a sweep may have long since eroded. Another explanation may be that selection was too weak relative to recombination to produce a sweep signature (Kaplan *et al.* 1988).

Interestingly, substitutions at positions 111 and 122 can be highly context dependent (i.e. epistatic) and N122D has a much smaller effect on ouabain sensitivity in the absence of a second substitution at position 111 (Price *et al.* 1990; Coppi *et al.* 1999). This implies that L111V may contribute to reducing ouabain sensitivity when coupled with N122H in monarchs. It also implies that the L111V substitution in other members of the *Danaus* genus may have less of an effect on ouabain sensitivity, because these species retain the ancestral asparagine at position 122. Future functional studies should test these hypotheses directly by investigating the individual and combined effects of both Q111L and L111V with N122H.

The possibility that observed changes at position 111 may have little effect on ouabain sensitivity on their own raises the interesting question of how other *Danaus* spp., which lack N122H, also manage to sequester CGs, albeit at generally lower levels relative to monarchs (Brower *et al.* 1975, 1978; Cohen 1985; Mebs *et al.* 2005). It may be that L111V has an effect on sensitivity to CGs in the context of other, as of yet uncharacterized, substitutions in the protein. Conversely, it may also be that mutations at one or more of many other sites in the protein result in an ouabain-resistant ATPα (Croyle *et al.* 1997), and substitutions at position 111 are not important for facilitating cardenolide resistance or sequestration. It is also important to consider that there is a great deal of variability in CG content and composition among populations and species of host plants, which may contribute to observed differences in CG content among *Danaus* spp. (Brower *et al.* 1972; Moranz & Brower 1998). Additional work is needed to better understand how substitutions to the H1–H2 extracellular domain of ATPα have influenced the evolution of host use in the Danainae butterflies. In particular, additional sampling of other members of this subfamily would allow us to better understand the ages of key amino acid substitutions and how the radiation of Danainae progressed in relation to these.

Functional mutagenesis and structural studies have identified numerous sites throughout ATPα at which substitutions could reduce the enzyme's sensitivity to cardenolides (Croyle *et al.* 1997; Ogawa *et al.* 2009; Yatime *et al.* 2011). Given this large number of potential targets, it is interesting that the same progression of three substitutions has occurred independently in the subfamilies Danainae and Eumolpinae. It is possible that nonsynonymous substitutions that could result in a cardenolide-resistant ATPα are limited and that the parallel changes observed reflect epistatic constraints on the evolution of this adaptation (Weinreich *et al.* 2006; Camps 2007).

Furthermore, much attention has focused on the H1–H2 extracellular domain of ATPα. Patterns of evolution at other potential targets for modulating cardenolide sensitivity scattered throughout the protein have not been documented in cardenolide-resistant insect taxa, even though such sites could also have importance in conferring resistance and an ability to sequester CG compounds (Palasis *et al.* 1996; Croyle *et al.* 1997). Other mechanisms to avoid the toxic effects of cardenolide consumption are also likely to be important in some herbivorous insects that feed on CG-rich plant material (Vaughan & Jungreis 1977; Torrie *et al.* 2004; Petschenka & Dobler 2009). To better understand cardenolide sequestration and the molecular changes that have contributed to this ability, the complete protein of other cardenolide-resistant forms of ATPα in other taxa (e.g. toads [*Bufo* spp.], fireflies [*Photinus* spp.] and *Hydra vulgaris*) could also be examined. Such information would allow us to better understand the evolution of cardenolide resistance and sequestration as well as host-plant-related adaptation and radiation in herbivorous insects.

Acknowledgements

M. Boppré, M. Schumer, W. Watt and two anonymous reviewers provided helpful comments on earlier drafts of this work. Thanks to A. Briscoe and L. Gilbert for providing queen butterflies. M. Scriber, M. Cheng and B. Sonke assisted with monarch collection. M. Boppré and D. Mebs provided *D. chrysippus*, *T. petiverana*, *A. ochlea*, *E. corinna* and *E. phaenareta* samples. This work was supported by NSF grant DEB-0946398.

References

Ackery PR (1988) Host plants and classification: a review of the nymphalid butterflies. *Biological Journal of the Linnean Society*, **33**, 95–203.

Ackery PR, Vane-Wright RI (1985) Patterns of plant utilization by danaine butterflies. In: *Proceedings of the 3rd Congress of European Lepidoterology* (ed. Heath J), pp. 3–6. Societas Europaea Lepidopterologica, Karlsruhe, Germany.

Agrawal AA, Fishbein M, Halitschke R, Hastings AP, Rabosky DL, Rasmann S (2009) Evidence for adaptive radiation from a phylogenetic study of plant defenses. *Proceedings of the National Academy of Sciences, USA*, **106**, 18067–18072.

Berenbaum MR, Favret C, Schuler MA (1996) On defining ''key innovations'' in an adaptive radiation: cytochrome P450s and Papilionidae. *The American Naturalist*, **148**, S139–S155.

Brower JV (1958) Experimental studies of mimicry in some North American butterflies: Part I. The monarch, *Danaus plexippus*, and viceroy, *Limenitis archippus archippus*. *Evolution*, **12**, 32–47.

Brower LP, Glazier SC (1975) Localizations of heart poisons in the monarch butterfly. *Science*, **188**, 19–25.

Brower LP, McEvoy PB, Williamson KL, Flannery MA (1972) Variation in cardiac glycoside content of monarch butterflies from natural populations in eastern North America. *Science*, **177**, 426–429.

Brower LP, Edmunds M, Moffit CM (1975) Cardenolide content and palatability of a population of *Danaus chrysippus* butterflies from West Africa. *Journal of Entomology Series A – Physiology & Behavior*, **49**, 183–196.

Brower LP, Gibson DO, Moffitt CM, Panchen AL (1978) Cardenolide content of *Danaus chrysippus* butterflies from three areas of East Africa. *Biological Journal of the Linnean Society*, **10**, 251–273.

Brower AVZ, Wahlberg N, Ogawa JR, Boppré M, Vane-Wright RI (2010) Phylogenetic relationships among genera of danaine butterflies (Lepidoptera: Nymphalidae) as implied by morphology and DNA sequences. *Systematics and Biodiversity*, **8**, 75–89.

Cameron SL, Lambkin CL, Barker SC, Whiting MF (2007) A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad

timescales with high precision. *Systematic Entomology*, **32**, 40–59.

Camps M (2007) Genetic constraints on protein evolution. *Critical Reviews in Biochemistry and Molecular Biology*, **42**, 313–326.

Canfield V, Rettig Emanuel J, Spickofsky N, Levenson R, Margolskee RF (1990) Ouabain-resistant mutants of the rat Na,K-ATPase α2 isoform identified by using an episomal expression vector. *Molecular and Cellular Biology*, **10**, 1367–1372.

Canfield VA, Xu KY, D'Aquila T, Shyjan AW, Levenson R (1992) Molecular cloning and characterization of the Na,K-ATPase from *Hydra vulgaris*: implications for enzyme evolution and ouabain sensitivity. *New Biologist*, **4**, 339–348.

Cohen JA (1985) Differences and similarities in cardenolide contents of queen and monarch butterflies in Florida and their ecological and evolutionary implications. *Journal of Chemical Ecology*, **11**, 85–103.

Cohen JA, Brower LP (1983) Cardenolide sequestration by the dogbane tiger moth (*Cycniatenera*; Arctiidae). *Journal of Chemical Ecology*, **9**, 521–532.

Coppi MV, Compton LA, Guidotti G (1999) Isoform-specific effects of charged residues at borders of the M1–M2 loop of the Na,K-ATPase α subunit. *Biochemistry*, **38**, 2494–2505.

Croyle ML, Woo AL, Lingrel JB (1997) Extensive random mutagenesis analysis of the Na^+/K^+-ATPase α subunit indentifies known and previously unidentified amino acid residues that alter ouabain sensitivity, Implications for ouabain binding. *European Journal of Biochemistry*, **248**, 488–495.

Dobler S (2004) The evolution of adaptations to plant secondary compounds in *Chrysochus* leaf beetles (Chrsomelidae, Eumolpinae). In: *New Developments in the Biology of Chrysomelidae* (eds Jolivet P, Santiago-Blay JA and Schmitt M), pp. 117–123. SPB Academic Publishing bv, The Hague, The Netherlands.

Dobler S, Farrell BD (1999) Host use evolution *Chrysochus* milkweed beetles: evidence from behavior, population genetics and phylogeny. *Molecular Ecology*, **8**, 1297–1307.

Dobler S, Daloze D, Pasteels JM (1998) Sequestration of plant compounds in a leaf beetle's secretion: cardenolides in *Chrysochus*. *Chemoecology*, **8**, 111–118.

Duffey SS, Scudder GGE (1972) Cardiac glycosides in North American Asclepiadaceae, a basis for unpalatability in brightly coloured Hemiptera and Coleoptera. *Journal of Insect Physiology*, **18**, 63–78.

Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. *Evolution*, **18**, 586–608.

Farrell BD, Mitter C (1994) Adaptive radiation in insects and plants: time and opportunity. *American Zoologist*, **34**, 57–69.

Farrell BD, Mitter C (1998) The timing of insect/plant diversification: might *Tetraopes* (Coleoptera: Cerambycidae) and *Asclepias* (Asclepiadaceae) have co-evolved? *Biological Journal of the Linnean Society*, **63**, 553–577.

Flook PK, Rowell CHF (1998) Inferences about orthopteroid phylogeny and molecular evolution from small subunit nuclear ribosomal DNA sequences. *Insect Molecular Biology*, **7**, 163–178.

Frick C, Wink M (1995) Uptake and sequestration of ouabain and other cardiac glycosides in *Danaus plexippus* (Lepidoptera: Danainae): evidence for a carrier-mediated process. *Journal of Chemical Ecology*, **21**, 557–575.

Futyuma D, Keese MC (1992) Evolution and coevolution of plants and phytophagous arthropods. In: *Herbivores: Their Interactions with Secondary Plant Metabolites*, 2nd edn (eds Rosenthal GA and Berenbaum MR), pp. 439–475. Academic Press, New York.

Holzinger F, Wink M (1996) Mediation of ardiac glycoside insensitivity in the monarch butterfly (*Danaus plexippus*): role of an amino acid substitution in the ouabain binding site of Na^+,K^+-ATPase. *Journal of Chemical Ecology*, **22**, 1921–1935.

Holzinger F, Frick C, Wink M (1992) Molecular basis for the insensitivity of the Monarch (*Danaus plexippus*) to cardic glycosides. *FEBS*, **3**, 477–480.

Hudson RR (2002) Generating samples under a Wright–Fisher neutral model. *Bioinformatics*, **18**, 337–338.

Hunt T, Bergsten J, Levkanicova Z *et al.* (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913–1916.

Jaisser F, Canessa CM, Horisberger JD, Rossier BC (1992) Primary sequence and functional expression of a novel ouabain-resistant Na,K-ATPase. *The Journal of Biological Chemistry*, **267**, 16895–16903.

Jewell EA, Lingrel JB (1992) Chimeric rat Na,K-ATPase α_1/α_3* isoforms: analysis of the structural basis for differences in Na^+ requirements in the α_1 and α_3* isoforms. *Annals of the New York Academy of Sciences*, **671**, 120–133.

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: *Mammalian Protein Metabolism III* (ed. Munro HN), pp. 21–132. Academic Press, New York.

Kaplan N, Darden T, Hudson R (1988) The coalescent process in models with selection. *Genetics*, **120**, 819–829.

Kergoat GK, Delobel A, Fédière G, Rü BL, Silvain J-F (2005) Both host-plant phylogeny and chemistry have shaped the African seed-beetle radiation. *Molecular Phylogenetics and Evolution*, **35**, 602–611.

Labeyrie E, Dobler S (2004) Molecular adaptation of *Chrysochus* leaf beetles to toxic compounds in their food plants. *Molecular Biology and Evolution*, **21**, 218–221.

Lewin B (2004) *Genes VIII*. Person Education, Inc., Upper Saddle River, New Jersey.

Lushai G, Smith DAS, Gordon IJ, Allen JA, Maclean N (2003) Incomplete sexual isolation in sympatry between subspecies of the butterfly *Danaus chrysippus* (L.) and the creation of a hybrid zone. *Heredity*, **90**, 236–246.

Mebs D, Zehner R, Schneider M (2000) Molecular studies on the ouabain binding site of the Na+,K+-ATPase in milkweed butterflies. *Chemoecology*, **10**, 201–203.

Mebs D, Reuss E, Schneider M (2005) Studies on the cardenolide sequestration in African milkweed butterflies (Danainae). *Toxicon*, **45**, 581–584.

Moranz R, Brower LP (1998) Geographic and temporal variation of cardenolide-based chemical defenses of queen butterflies in Northern Florida. *Journal of Chemical Ecology*, **24**, 905–932.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nishida R (2002) Sequestration of defensive substances from plants by Lepidoptera. *Annual Review of Entomology*, **47**, 57–92.

Ogawa H, Shinoda T, Cornelius F, Toyoshima C (2009) Crystal structure of the sodium-potassium pump (Na^+,K^+-ATPase) with bound potassium and ouabain. *Proceedings of the National Academy of Sciences, USA*, **106**, 13742–13747.

Palasis M, Kuntzweiler TA, Argüello JM, Lingrel JB (1996) Ouabain interactions with the H5–H6 hairpin of the Na,K-ATPase reveal a possible inhibition mechanism via the cation binding domain. *The Journal of Biological Chemistry*, **271**, 14176–14182.

Petschenka G, Dobler S (2009) Target-site sensitivity in a specialized herbivore towards major toxic compounds of its host plant: the Na^+K^+-ATPase of the oleander hawk moth (*Daphnis nerii*) is highly susceptible to cardenolides. *Chemoecology*, **19**, 235–239.

Price EM, Lingrel JB (1988) Structure-function relationships in the Na,K-ATPase alpha-subunit: site-directed mutagenesis of glutamine-111 to arginine and asparagine-122 to aspartic acid generates a ouabain-resistant enzyme. *Biochemistry*, **27**, 8400–8408.

Price EM, Rice DA, Lingrel JB (1989) Sire-directed mutagenesis of a conserved, extracellular aspartic-acid residue affects the ouabain sensitivity of sheep Na,K-ATPase. *Journal of Biological Chemistry*, **264**, 21902–21906.

Price EM, Rice DA, Lingrel JB (1990) Structure-function studies of Na,K-ATPase. *The Journal of Biological Chemistry*, **265**, 6638–6641.

Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **160**, 1179–1189.

Regier JC, Fang QQ, Mitter C, Peigler RS, Friedlander TP, Alma Solis M (1998) Evolution and phylogenetic utility of the *period* gene in Lepidoptera. *Molecular Biology and Evolution*, **15**, 1172–1182.

Reidenbach KR, Cook S, Bertone MA, Harbach RE, Wiegmann BM, Besansky NJ (2009) Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evolutionary Biology*, **9**, 298.

Ritland DB, Brower LP (1993) A reassessment of the mimicry relationships among viceroys, queens, and monarchs in Florida. In: *Biology and Conservation of the Monarch Butterfly* (eds Malcolm SB and Zalucki MP), pp. 129–139. Science Series No. 38, Natural History Museum of Los Angeles County, CA, USA.

Schatzmann HJ (1964) The role of Na^+ and K^+ in the ouabain-inhibition of the Na^+ + K^+-activated membrane adenosine triphosphatase. *Biochimicaet Biophysica Acta*, **94**, 89–96.

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

Torrie LS, Radford JC, Southall TD, Kean L, Dinsmore AJ, Davies SA (2004) Resolution of the insect ouabain paradox. *Proceedings of the National Academy of Sciences*, **101**, 13689–13693.

Vaughan GL, Jungreis AM (1977) Insensitivity of lepidopteran tissues to ouabain: physiological mechanisms for protection from cardiac glycosides. *Journal of Insect Physiology*, **23**, 585–589.

Wahlberg N, Leneveu J, Kodandaramaiah U *et al.* (2009) Nymphalid butterflies diversity following near demise at the Cretaceous/Tertiary boundary. *Proceedings of the Royal Society, B*, **276**, 4295–4302.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, **312**, 111–114.

Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant-insect coevolutionary key innovation. *Proceedings of the National Academy of Sciences*, **104**, 20427–20431.

Wiegmann BM, Trautwein MD, Kim JW *et al.* (2009) Single-copy nuclear genes resolves the phylogeny of the holometabolous insects. *BMC Biology*, **7**, 34.

Yatime L, Laursen M, Morth JP, Esmann M, Nissen P, Fedosova NU (2011) Structural insights into the high affinity binding of cardiotonic steroids to the Na^+,K^+-ATPase. *Journal of Structural Biology*, **174**, 296–306.

Zhang M, Cao T, Jin K *et al.* (2008) Estimating divergence times among subfamilies in Nymphalidae. *Chinese Science Bulletin*, **53**, 2652–2658.

Zhu H, Casselman A, Reppert SM (2008) Chasing migration genes: a brain expressed sequences tag resource for summer and migratory monarch butterflies (*Danaus plexippus*). *PLoS ONE*, **3**, e1345.

M.L.A. is interested in the evolution of reproductive isolation between ecologically divergent populations. Y.Z. is interested in the adaptive evolution of proteins and non-coding DNA between species and among divergent natural populations. P.A. is interested in using genetic/genomic technological advances and population genetic methodologies to elucidate the genetic basis of specific adaptations and quantify the extent and mode of natural selection in the genome.

Data accessibility

cDNA sequences for monarch and queen: GenBank accessions nos: JN846931 and JN846930.

The raw data are provided as Appendices S1–S4 (Supporting information).

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Partial sequences of ATPα from 69 surveyed monarch butterflies in FASTA format. The intron used in analysis (see Methods) begins at position 100 and ends at position 188.

Appendix S2 Partial sequences of ATPα from seven surveyed queen butterflies in FASTA format. The intron used in analysis (see Methods) begins at position 46 and ends at position 126.

Appendix S3 Partial sequences for ATPα from additionally surveyed Nymphalidae butterflies in FASTA format (see Table S1).

Appendix S4 Sequence for the H1–H2 associated intron from additionally surveyed Nymphalidae butterflies in FASTA format. This intron is located between positions 45 and 46 of the coding sequence (Appendix S3).

Table S1 Detailed information for the species used in our population genetic and phylogenetic analyses of the Danainae subfamily.

Table S2 Summary of polymorphic and divergent sites in the intron of the H1–H2 extracellular domain of Na,K-ATPase.

Table S3 Taxa used in our phylogenetic study.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.