

Sex Chromosome Mosaicism and Hybrid Speciation among Tiger Swallowtail Butterflies

Krushnamegh Kunte^{1,2*}, Cristina Shea², Matthew L. Aardema³, J. Mark Scriber⁴, Thomas E. Juenger¹, Lawrence E. Gilbert¹, Marcus R. Kronforst^{2*}

1 Section of Integrative Biology, University of Texas, Austin, Texas, United States of America, **2** FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts, United States of America, **3** Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, United States of America, **4** Department of Entomology, Michigan State University, East Lansing, Michigan, United States of America

Abstract

Hybrid speciation, or the formation of a daughter species due to interbreeding between two parental species, is a potentially important means of diversification, because it generates new forms from existing variation. However, factors responsible for the origin and maintenance of hybrid species are largely unknown. Here we show that the North American butterfly *Papilio appalachiensis* is a hybrid species, with genomic admixture from *Papilio glaucus* and *Papilio canadensis*. *Papilio appalachiensis* has a mosaic phenotype, which is hypothesized to be the result of combining sex-linked traits from *P. glaucus* and *P. canadensis*. We show that *P. appalachiensis*' Z-linked genes associated with a cooler thermal habitat were inherited from *P. canadensis*, whereas its W-linked mimicry and mitochondrial DNA were inherited from *P. glaucus*. Furthermore, genome-wide AFLP markers showed nearly equal contributions from each parental species in the origin of *P. appalachiensis*, indicating that it formed from a burst of hybridization between the parental species, with little subsequent backcrossing. However, analyses of genetic differentiation, clustering, and polymorphism based on molecular data also showed that *P. appalachiensis* is genetically distinct from both parental species. Population genetic simulations revealed *P. appalachiensis* to be much younger than the parental species, with unidirectional gene flow from *P. glaucus* and *P. canadensis* into *P. appalachiensis*. Finally, phylogenetic analyses, combined with ancestral state reconstruction, showed that the two traits that define *P. appalachiensis*' mosaic phenotype, obligatory pupal diapause and mimicry, evolved uniquely in *P. canadensis* and *P. glaucus*, respectively, and were then recombined through hybridization to form *P. appalachiensis*. These results suggest that natural selection and sex-linked traits may have played an important role in the origin and maintenance of *P. appalachiensis* as a hybrid species. In particular, ecological barriers associated with a steep thermal cline appear to maintain the distinct, mosaic genome of *P. appalachiensis* despite contact and occasional hybridization with both parental species.

Citation: Kunte K, Shea C, Aardema ML, Scriber JM, Juenger TE, et al. (2011) Sex Chromosome Mosaicism and Hybrid Speciation among Tiger Swallowtail Butterflies. *PLoS Genet* 7(9): e1002274. doi:10.1371/journal.pgen.1002274

Editor: Nancy A. Moran, Yale University, United States of America

Received: April 7, 2011; **Accepted:** July 13, 2011; **Published:** September 8, 2011

Copyright: © 2011 Kunte et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by University of Texas fellowships and research grants to KK and Faculty Research Grants to LEG and TEJ; NSF grant DEB-0716683 to JMS; NSF Award DEB-0546316 to TEJ; NSF grant DEB-0640512 to LEG; and NIH NIGMS Grant GM068763, NSF grant DEB-1020355, and Harvard University Bauer Fellow funds to MRK. Additional infrastructural and financial support was provided by the Michigan State Experiment Station Project # 01644 to JMS and College of Natural Sciences, Plant Science Fellowship, a Scriber Scholars in Butterfly Biology and Conservation Award (Dept. Entomology, MSU), and an NSF REU Award (DEB-0821958 to JMS) to MLA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: KKunte@cgr.harvard.edu (KK); mkronforst@cgr.harvard.edu (MRK)

Introduction

Inter-specific hybridization is widespread in nature and may have important consequences in evolution, from the transfer of adaptive alleles between species to the formation of hybrid species [1–8]. Although prevalent in plants, hybrid speciation is apparently uncommon among animals, which may be explained by two factors: (a) hybrid populations may be weakly reproductively isolated from the parental species, so instead of maintaining their genomic identity, they may fuse with one of the parental species by backcrossing, and (b) hybrids may be unable to compete for ecological niches already occupied by parental species [3–5]. Hence, hybrid populations are likely to evolve as a species only when hybrids mate assortatively with other hybrids [7], and/or when they adapt to a new environment [6]. Both processes involve selection on specific phenotypic and ecological traits that favor the

origin and maintenance of hybrid species. Selective introgression of genes responsible for divergent ecological traits from both parental species into a hybrid species would provide strong evidence for natural selection in the origin of a hybrid species [9]. However, selective regimes that may favor the formation of hybrid species are poorly understood and genetic evidence for selective introgression of ecologically important traits in a hybrid species is limited [9,10]. For instance, several putative examples of animal hybrid species are known, ranging from fish to crustaceans and insects [11], but the potential role of natural selection in their origin and maintenance is largely unknown. Also, there are no examples of hybrid species that maintain their genomic identity while in contact with both parental species, which is important because sympatry may suggest that hybrid species are maintained by natural selection. Indeed, other known or suspected hybrid species are allopatric relative to either one or both parental species

Author Summary

Hybridization between two species may sometimes lead to the formation of a third species. While this phenomenon is of interest because it can generate new species from existing variation, it is rare in animals. Therefore, the factors responsible for the origin and maintenance of hybrid species are largely unknown. Here we show that the butterfly *Papilio appalachiensis* appears to be a hybrid species, which received nearly equal contributions to its genome from *P. glaucus* and *P. canadensis*. Two sets of traits define *P. appalachiensis*' hybrid phenotype: like *P. canadensis* it inhabits a cold habitat and has a single generation every year, and like *P. glaucus* it mimics a toxic butterfly and its females are dimorphic. The genes responsible for these traits are on two different sex chromosomes. Our genetic data show that *P. appalachiensis* inherited the sex chromosome associated with the cold habitat from *P. canadensis*, whereas it inherited the sex chromosome associated with mimicry and dimorphism from *P. glaucus*. Historical climate change may have facilitated the origin of *P. appalachiensis*, which our data suggest coincided with the last interglacial period in North America. This example shows the importance of sex-linked traits and a potential role for natural selection in the origin and maintenance of hybrid species.

[6,7,11]. Here we show that the tiger swallowtail butterfly *Papilio appalachiensis* exhibits many hallmarks of a hybrid species, including a mosaic genome derived from *P. glaucus* and *P. canadensis*. At the same time, its genome is significantly differentiated while being in contact with both its parental species. We further show that its hybrid phenotype, which appears to be under selection, was produced by combining sex chromosomes of its parental species. Thus, *Papilio appalachiensis* may be an informative case for understanding the role of natural selection in both the origin and maintenance of hybrid species.

Tiger swallowtail butterflies of North America (Figure 1A) are a monophyletic group consisting of eight closely-related species. Sister species *glaucus* and *canadensis* show clear and strong evidence for interspecific divergence and multiple forms of reproductive isolation: one-way assortative mate preference [12,13], reduced hatching success of hybrid eggs [13,14], Haldane's Rule and hybrid incompatibility in *glaucus/canadensis* crosses [14,15], divergent thermal habitat preference [16–18], differential host-plant preference/usage [18–22] and larval development times [23], and differential survival on preferred host plants [14,19,24] (also see Materials and Methods). This broad range of isolating mechanisms shows that *glaucus* and *canadensis* are good biological species [13,25]. They have diverged ecologically along a steep thermal gradient in spite of hybridizing across a long and narrow hybrid zone [25–27]. This hybrid zone is characteristically bimodal, where hybrids are uncommon relative to the parental species [28] (also see Materials and Methods). The hybrid zone is maintained by thermal ecology: *glaucus* is adapted to a warmer thermal habitat compared to *canadensis*. This has shaped key ecological differences between the two species, including voltinism and pupal diapause (Figure 1B). These traits are Z-linked (Lepidoptera have ZZ males and ZW females) [15,21,29]. The thermal landscape also contributes indirectly to the evolution of Batesian mimicry, which is under frequency-dependent selection and produces sexual dimorphism and female dimorphism in *glaucus* [30,31]. *Papilio glaucus* is palatable to vertebrate predators, and has two female forms: a yellow, male-like, non-mimetic form and a melanic form that mimics the chemically defended *Battus philenor* (Figure 1A)

[32]. The distribution of *B. philenor* is limited in the north by its larval host plant, producing a latitudinal gradient in the frequency of the *glaucus* mimetic female form [30,33,34]. The Mendelian locus controlling female color in *glaucus* is W-linked, with a Z-linked mimicry enabler allele nearly fixed in *glaucus* and a mimicry suppressor allele fixed in *canadensis* [26,29].

Recently described *P. appalachiensis* appears to be a phenotypic mixture of *glaucus* and *canadensis*, with which it is sympatric and parapatric, respectively (Figure 1A), but is apparently reproductively isolated from both (see Materials and Methods, and below). It shows affinity with *canadensis* for traits related to its cooler thermal habitat but with *glaucus* for traits related to mimicry and sexual/female dimorphism (Figure 1B) [35–38]. This mosaic phenotype, which is critical to its existence in cooler, high-elevation habitats along the Appalachian Mountains but within the range of *B. philenor*, has led to the suggestion that *appalachiensis* may be a hybrid species [35,37]. In support of this hypothesis, preliminary allozyme data have shown that *appalachiensis* has a *canadensis*-like allele at the Z-linked *Lactate Dehydrogenase (Ldh)* gene and a *glaucus*-like allele at the Z-linked *Phosphogluconate Dehydrogenase (Pgd)* gene [38].

Here we test the hypothesis that *appalachiensis* is a hybrid species with genomic admixture from *glaucus* and *canadensis*. If it is a hybrid species, its specific mixture of traits combined with the known sex-linkage of these traits make a clear prediction regarding the ancestry of *appalachiensis*' sex chromosomes: *appalachiensis* has inherited most of its Z chromosome from *canadensis* and its W chromosome from *glaucus*. We also test whether it has genome-wide admixture from the parental species, as expected of hybrid species that are formed with little subsequent backcrossing with parental species. In addition, we test alternative hypotheses regarding the apparently mixed genotype and phenotype of *appalachiensis*. Finally, we estimate divergence times between *glaucus*, *canadensis* and *appalachiensis* and study the evolutionary history of the traits that make up *appalachiensis*' hybrid phenotype in order to shed light on the evolutionary dynamics of the origin and maintenance of this hybrid species.

Results/Discussion

We tested the prediction that *appalachiensis* has mosaic sex chromosomes by sequencing six genes along the length of the Z chromosome and the mitochondrial *cytochrome oxidase I (COI)* gene (Materials and Methods, Table S2). While there are currently no known W-linked molecular markers in butterflies, we can infer the history of the W chromosome by sequencing mitochondrial DNA because both are maternally inherited and hence linked in Lepidoptera (see Materials and Methods). As predicted, our sequence data showed clear genetic discordance with *appalachiensis* having a mitochondrial genome similar to that of *glaucus*, and a Z chromosome similar to *canadensis* (Figure 2, Tables S3 and S4).

While the observed sex chromosome mosaicism suggests that *appalachiensis* evolved by combining key, sex-linked ecological traits from *glaucus* and *canadensis*, it does not tell us whether *appalachiensis* has a hybrid ancestry across the rest of its genome. Such widespread genomic admixture would be expected for a hybrid species formed with little subsequent backcrossing with the parental species. Alternatively, *appalachiensis* may be largely similar to one parental species across its genome, with a selectively introgressed W (and mtDNA) or Z chromosome from the other parental species [36]. To test these alternatives, we screened 2,035 nuclear amplified fragment length polymorphism (AFLP) markers and identified 249 that distinguished *glaucus* from *canadensis*. We then genotyped these ancestry informative markers in *glaucus*,

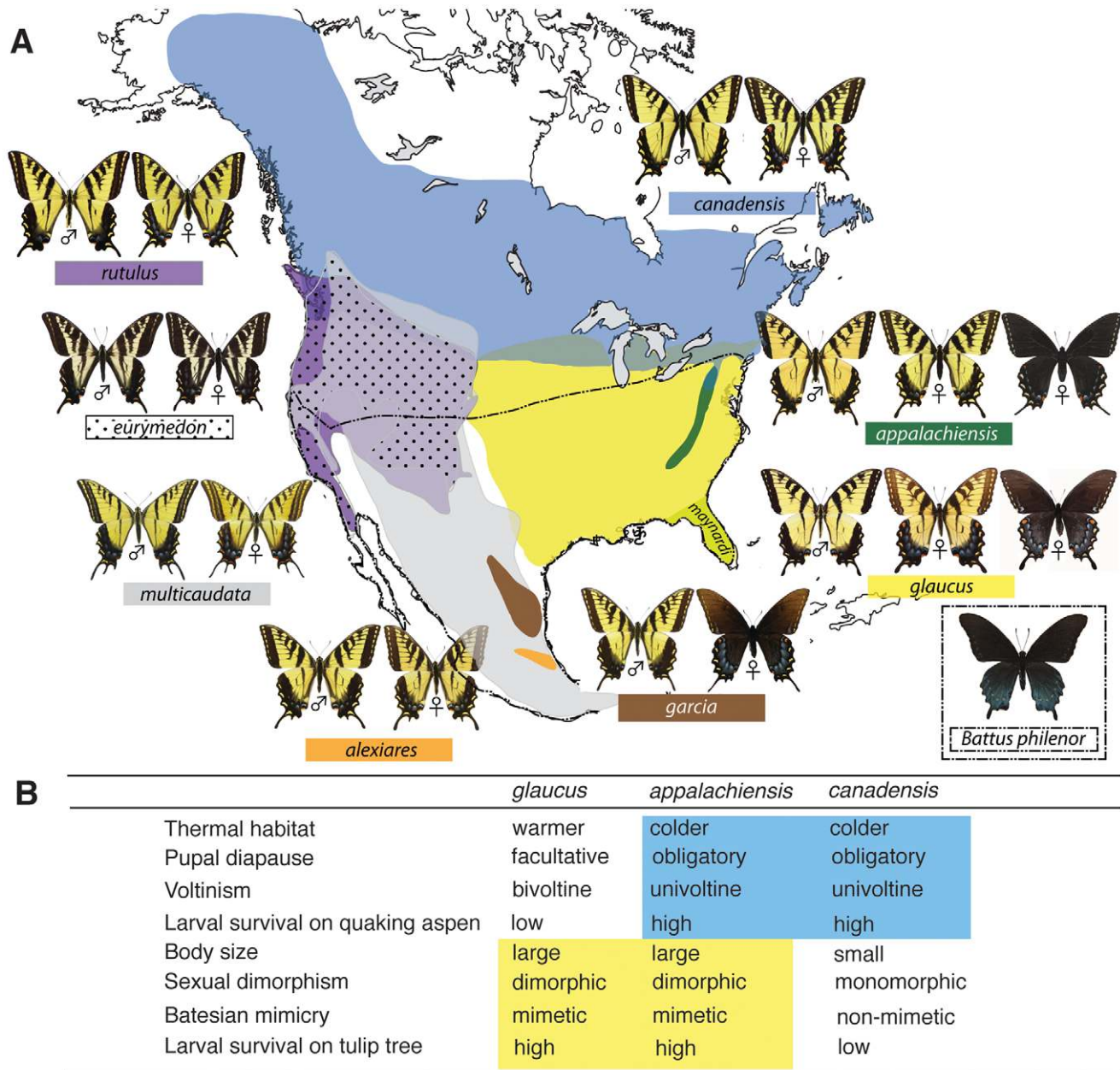


Figure 1. The distributional ranges and hybrid zones of tiger swallowtails, and the hybrid phenotype of *Papilio appalachiensis*. (A) *Papilio appalachiensis* is endemic to mid- and high elevations in the Appalachian Mountains and sympatric with *glaucus* throughout its range, but presumably parapatric with *canadensis* in its northernmost range [35,36] (see Materials and Methods). Also shown is the range of *Battus philenor*, Batesian model for the mimetic *glaucus*, *appalachiensis* and *garcia* melanic female forms. (B) Ecological and morphological differentiation between *glaucus* and *canadensis*, and their admixture in *appalachiensis* [35–38] (also see Figure S1). doi:10.1371/journal.pgen.1002274.g001

canadensis, *appalachiensis* and laboratory-generated *glaucus* x *canadensis* F1 and F2 hybrids. We used the program STRUCTURE to genetically cluster the samples under the assumption that they represented two, three or four admixed populations. This analysis revealed that *appalachiensis* has genome-wide admixture indicative of historical hybridization between *glaucus* and *canadensis* (Figure 3A, Figure S2). Under the assumption of two populations (K = 2), *appalachiensis* and laboratory-generated hybrids showed similar and nearly equal admixture from *glaucus* and *canadensis*. At K = 3, *appalachiensis* formed its own cluster, and at K = 4 it remained a distinct cluster while subdivision within *glaucus* became apparent.

In contrast, at K = 3 and K = 4, the laboratory-generated hybrids continued to show similar admixture from both parental species. A locus-by-locus analysis of molecular variance (AMOVA) of the AFLP dataset further confirmed widespread genomic admixture in *appalachiensis*: allele frequencies at 67 AFLP markers were similar to *glaucus* and significantly different from those of *canadensis*, 74 AFLP markers were similar to *canadensis* and significantly different from *glaucus*, and 92 markers had frequencies intermediate between *glaucus* and *canadensis* but significantly different from neither (Figure 3B). This analysis also revealed 16 AFLPs for which the frequency in *appalachiensis* was significantly different from both

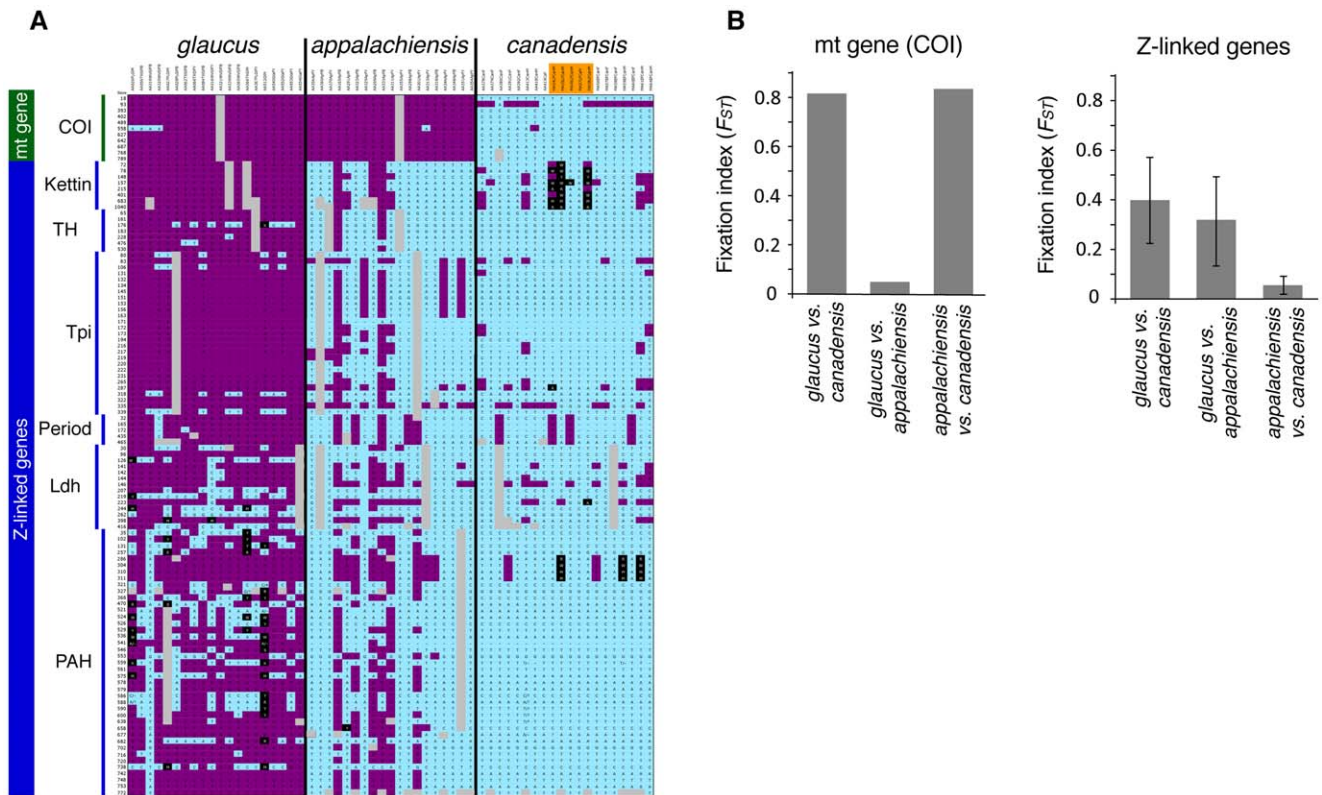


Figure 2. Genotypic differentiation between *glaucus* and *canadensis*, and the mismatch in mitochondrial and Z-linked genes in *appalachiensis*. (A) *appalachiensis* genotypes at loci that were significantly different ($p < 0.001$) between *glaucus* and *canadensis*, as judged by F_{ST} values from a locus-by-locus AMOVA comparing *glaucus* and *canadensis*. Genotypes are nucleotide bases at specific SNP or indel polymorphisms, which can be diploid (Z-linked polymorphisms scored in males) or haploid (mtDNA, and Z-linked polymorphisms scored in females). Color code: purple: genotypes characteristic of *glaucus*; light blue: genotypes characteristic of *canadensis*; black: heterozygotes; grey: missing data; orange: late flight *canadensis*. (B) Species pair-wise F_{ST} values for the mitochondrial and Z-linked genes (see Table S4 for individual values for each gene and species pair-wise comparisons). doi:10.1371/journal.pgen.1002274.g002

parental species (Figure 3B). Importantly, co-occurring *appalachiensis* and *glaucus* individuals, as identified morphologically at the time of collection from Spruce Knob in May and June 2006 (see Materials and Methods), showed species-specific sequences (Figure 2A) and AFLP signatures (Figure 3A). This suggests that these two co-occurring species are genetically distinct and that the clusters seen in Figure 2 and 3 are not merely geographically isolated subpopulations of the same species.

Examining the evolutionary context of the hybrid phenotype of *appalachiensis* requires a well-resolved phylogeny of tiger swallowtails. However, resolving this phylogeny from sequence data is challenging owing to recent divergence, incomplete lineage sorting and ongoing hybridization [39,40] (also see Materials and Methods). Therefore, we used 1,607 polymorphic AFLP markers to generate a bootstrap-supported neighbor-joining tree of the entire clade. All tiger swallowtail nodes were strongly supported, except the *glaucus*, *canadensis* and *appalachiensis* clades (Figure 4A; also see Table S5). Within the *glaucus*-*canadensis*-*appalachiensis* groups, individuals formed species-specific clusters with only 12 out of the 184 samples clustering outside their own species (Figure 4A). We resolved the ancestral nodes between *appalachiensis*, *glaucus* and *canadensis* based on Figure 2 and Figure 3, and mapped relevant ecological traits (Table S6) on the resulting phylogeny (Figure 4B). We found that the mimetic female form suppressor is basal and widespread among tiger swallowtails. Mimicry and the enabler either evolved independently in *garcia* and *glaucus*, or were

lost in *alexiares* and *canadensis*. The ancestral yellow female form was uniquely lost in *garcia*, whereas female dimorphism was uniquely gained in *glaucus* and then introgressed into *appalachiensis*. Similarly, obligatory pupal diapause and univoltinism evolved uniquely in *canadensis* and introgressed into *appalachiensis*. Thus, we not only identified the specific traits that recombined to form the hybrid species, we also traced the origin and spread of those traits through evolutionary history, down to the time they were brought together through hybridization to generate a new species.

The evolutionary history and direction of introgression in *appalachiensis* contrasts with the phenomenon of *canadensis* “late flight”, which has emerged recently near the northern limit of the *glaucus*-*canadensis* hybrid zone in Vermont [38,41]. The late flight occurs in July and is allochronic relative to true (“early flight”) *canadensis*, which flies in May and June [38,41]. The late flight has been hypothesized to be a result of hybridization between *glaucus* and *canadensis*, potentially representing an early stage in the evolution of an *appalachiensis*-like entity [41]. We used our DNA sequence and AFLP data to test whether late flight *canadensis* is of hybrid origin, and whether it is similar to *appalachiensis*. We found that unlike *appalachiensis*, the late flight is entirely *canadensis*-like at both its mitochondrial and Z-linked genes, except for three late flight individuals that were heterozygous for *glaucus*- and *canadensis*-like haplotypes at *Kettin* (Figure 2A). However, AFLP data show a strong signature of genome-wide admixture in late flight individuals (Figure 3C). Furthermore, the AFLP signature of the

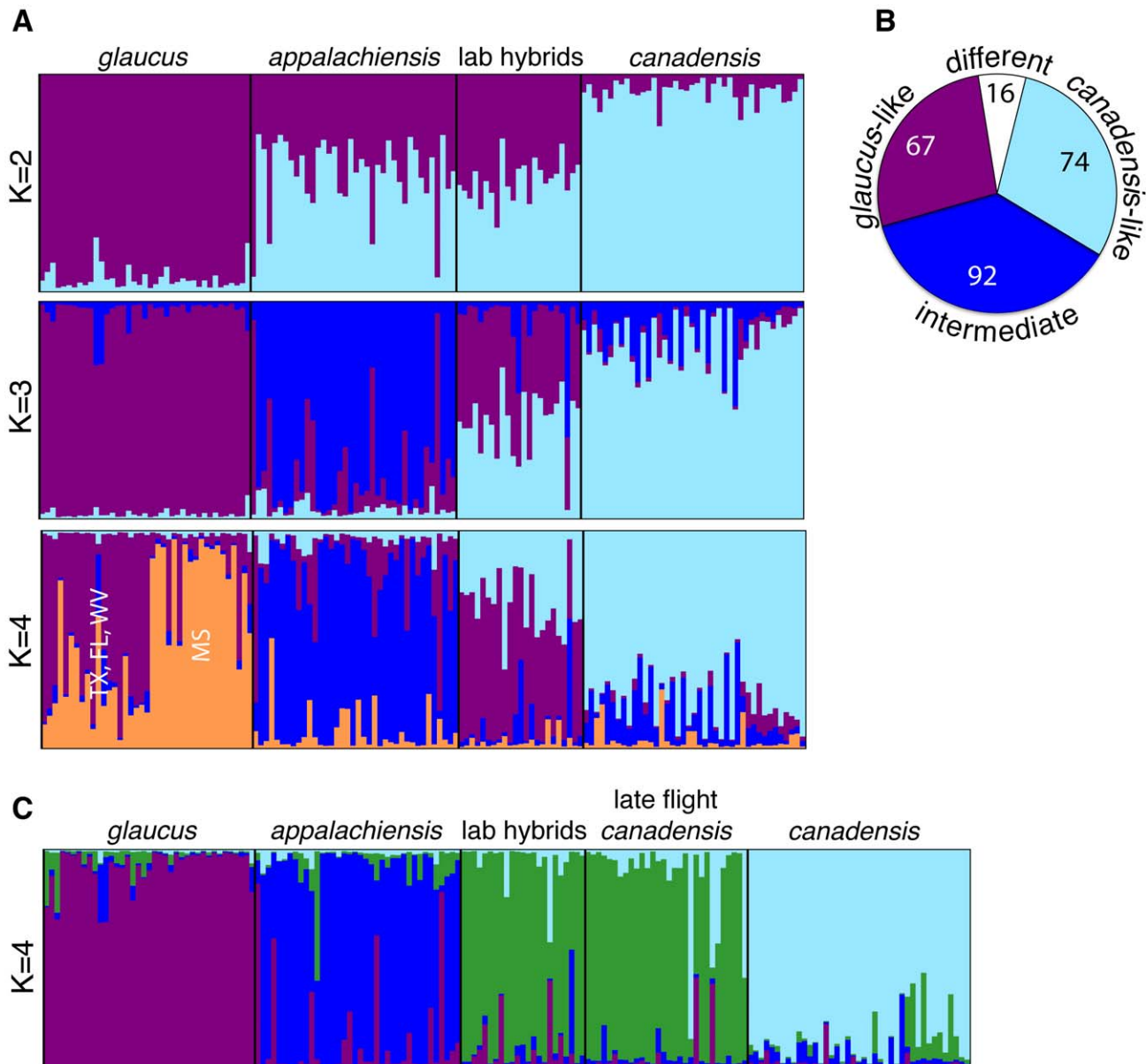


Figure 3. Genomic admixture in *appalachiensis* showing its hybrid origin and its contrast with laboratory-generated hybrids and late flight *canadensis*. (A) Population clustering of AFLP data in STRUCLURE under the assumption of two, three and four populations, comparing *appalachiensis* with laboratory-generated *glaucus* x *canadensis* hybrids. (B) *appalachiensis* AFLP allele frequencies with respect to *glaucus* and *canadensis*, based on species pair-wise locus-by-locus AMOVAs. Allele frequencies of “*glaucus*-like” AFLPs were significantly different from *canadensis*, “*canadensis*-like” AFLPs were significantly different from *glaucus*, “intermediate” were intermediate between *glaucus* and *canadensis* but significantly different from neither, and “different” were significantly different from both *glaucus* and *canadensis*. (C) Population clustering in STRUCLURE under the assumption of four populations, showing genomic similarity between the laboratory-generated hybrids and late flight *canadensis*, and distinctiveness of *appalachiensis* (also see Figure S2). For (A) and (C), admixture proportions of the sampled individuals, rather than their assignment probabilities, are shown.

doi:10.1371/journal.pgen.1002274.g003

late flight is indistinguishable from that of laboratory-generated *glaucus* x *canadensis* F1 and F2 hybrids but distinct from *appalachiensis* (Figure 3C, Figure S2B), as expected if the late flight is a result of recent hybridization. The recent history of hybridization in the late flight is also supported by coalescent simulations (Figure S4, also see below and in Materials and Methods). Together, these data show that its history of introgression is dissimilar to *appalachiensis*. The late flight has the potential to speciate through allochronic flight period and larval host plant specialization [20,41]; however, unlike *appalachiensis*, it may be a transient phase

in the northward movement of the *glaucus*-*canadensis* hybrid zone in a changing thermal landscape. Moreover, although selection may maintain some characteristics of a mosaic sex chromosome in the late flight [20,37], our data suggest that the late flight may not be experiencing natural selection for the same sex chromosome mosaicism that seems to have been essential in generating and maintaining the hybrid phenotype of *appalachiensis*.

We tested two alternative hypotheses for the apparent hybrid ancestry of *appalachiensis*: (1) *appalachiensis*, *glaucus* and *canadensis* diverged from one another in a standard bifurcating manner and

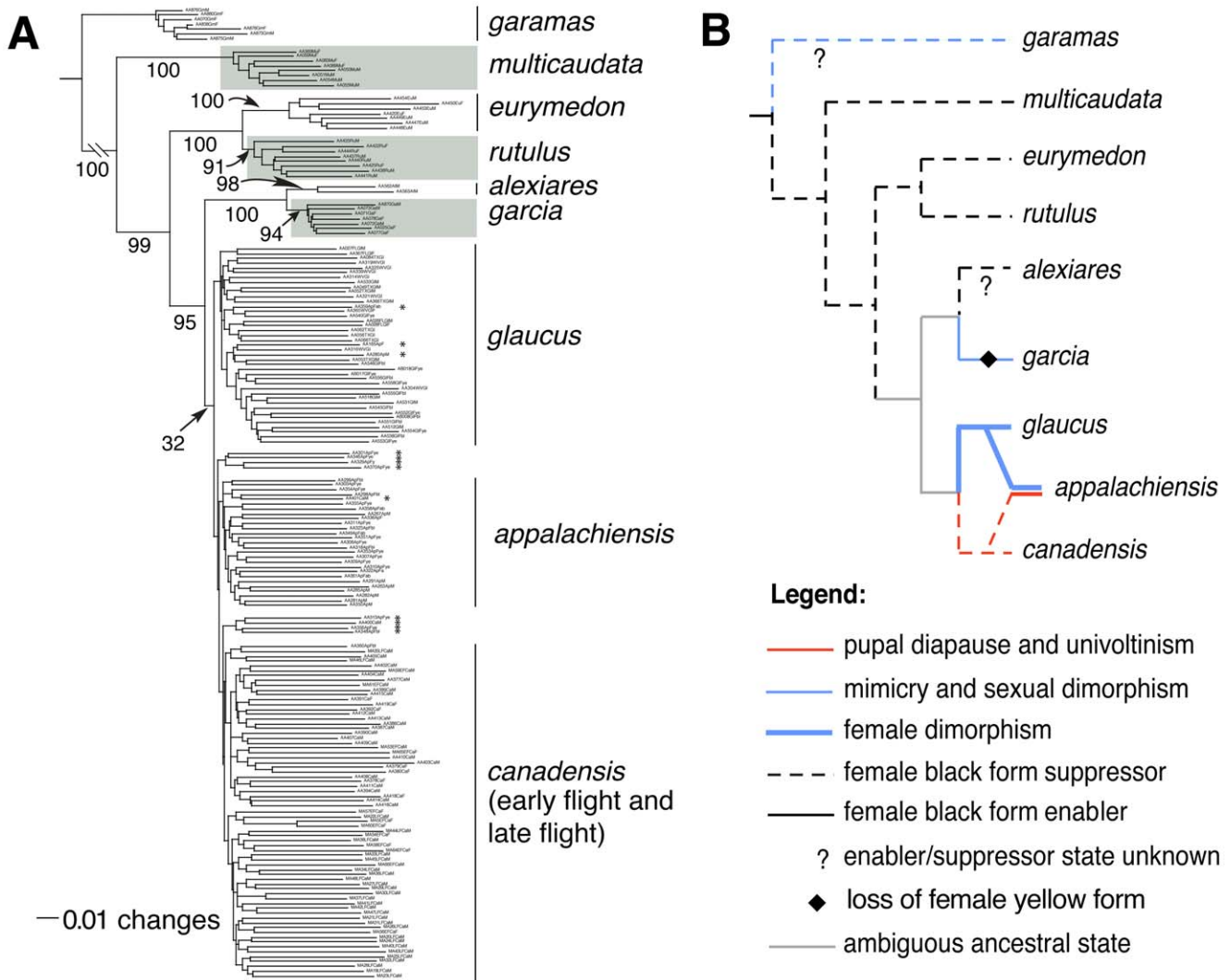


Figure 4. Phylogenetic relationships and character evolution among tiger swallowtails. (A) AFLP-based neighbor-joining tree, with percentage bootstrap support shown for branches. The ten *appalachiensis* and two *canadensis* samples that cluster outside their species are marked with asterisks. (B) Character evolution based on the AFLP phylogeny. doi:10.1371/journal.pgen.1002274.g004

ancestral variation sorted out in such a way that *appalachiensis* now appears to have a hybrid genotype and phenotype, and (2) *appalachiensis* is not a reproductively isolated species but a very recent hybrid population that is constantly supplied by ongoing hybridization between *glaucus* and *canadensis*. We addressed the first hypothesis with our DNA sequence data by estimating divergence times and rates of historical introgression among the three species with the coalescent approach implemented in IMA2. Instead of supporting a bifurcating model of speciation, this analysis showed that *glaucus* and *canadensis* diverged approximately 580,000 years ago, whereas *appalachiensis* diverged from both the parental species approximately 100,000 years ago during the Pleistocene (Figure 5). The IMA2 analysis also showed essentially unidirectional introgression from *glaucus* and *canadensis* into *appalachiensis* (Figure 6), consistent with the hypothesized hybrid origin of *appalachiensis*.

We tested the second hypothesis by comparing among the three focal species: (a) the extent of linkage-disequilibrium, and (b) the proportion of species-specific genetic polymorphisms and haplotypes. Recent hybridization is expected to produce elevated linkage disequilibrium as a result of bringing together two distinct

chromosomal copies from each parent species, the linkage disequilibrium decaying through successive generations as a result of recombination [42]. Indeed, linkage disequilibrium in both our DNA sequence and AFLP data was elevated in the laboratory-generated hybrids and late flight *canadensis*, both of which are products of very recent hybridization (Table 1). In contrast, linkage disequilibrium was significantly lower in *appalachiensis* and within the range of that seen in *glaucus* and *canadensis* (Table 1). Together, this evidence supports *appalachiensis*' historical origin and highlights the differences in the hybrid history between *appalachiensis*, late flight *canadensis* and laboratory-generated hybrids. If *appalachiensis* is indeed an old and subsequently isolated hybrid species, then it should possess unique mutations that it would have accumulated since it split from its parental species. These unique mutations should be detectable even if their number may be relatively small and they may not be fixed given: (a) the short estimated divergence time of ca 100,000 years, and (b) possibly low level of ongoing hybridization with the parental species. We calculated the proportion of species-specific single nucleotide polymorphisms and haplotypes among the three focal species

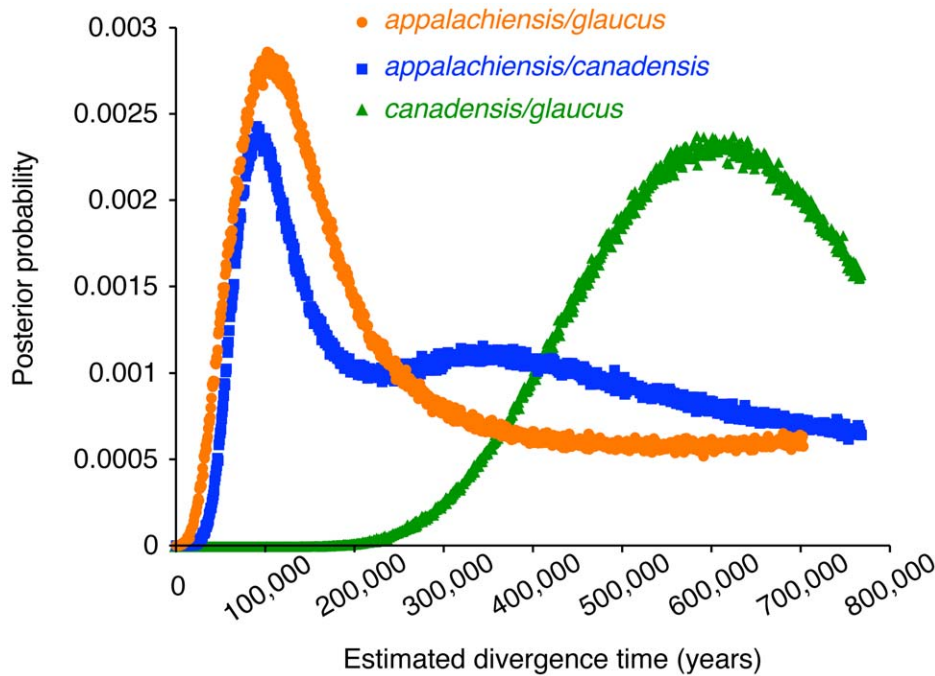


Figure 5. Estimated divergence times between the parental *glaucus* and *canadensis* and the hybrid *appalachiensis*. Dates of divergence estimated by IMa2 are: (a) *appalachiensis* and *glaucus*: ca 100,000 years ago, (b) *appalachiensis* and *canadensis*: ca 90,000 years ago, (c) *glaucus* and *canadensis*: ca 580,000 years ago. doi:10.1371/journal.pgen.1002274.g005

based on our sequence data from seven genes (Table 2). The number and proportion of species-specific polymorphisms among *glaucus*, *canadensis* and *appalachiensis* varied greatly among genes and species, but all three species had a substantial number of species-specific polymorphisms (Table 2). The proportion of species-specific polymorphisms averaged across the seven genes was 0.58 (± 0.221) for *canadensis*, 0.59 (± 0.158) for *glaucus*, and 0.41 (± 0.19) for *appalachiensis*. Thus, *appalachiensis* had lower average species-specific polymorphisms compared to its parental species, as expected from its younger age, but the difference was not significant (ANOVA, $F_{(2,18)} = 1.90$, $p = 0.179$). We conclude that *appalachiensis* appears to have been isolated from the parental species long enough to accumulate species-specific mutations

whose numbers are comparable with those of the parental species. This pattern is robust and holds true if we consider the number of species-specific haplotypes instead of polymorphisms: the proportion of species-specific haplotypes, averaged across all seven genes, was $0.71(\pm 0.270)$ for *canadensis*, $0.9(\pm 0.06)$ for *glaucus*, and $0.68(\pm 0.098)$ for *appalachiensis* (ANOVA: $F_{(2,18)} = 3.47$, $p = 0.0533$) (Table 2). These polymorphism and haplotype data support the hypothesis that *appalachiensis* is a good species rather than a recent or constantly supplied hybrid population. Moreover, additional lines of evidence are in conflict with the second hypothesis: (a) *appalachiensis* has a unique morphological and behavioral phenotype (Figure 1B and Figure S1), (b) there is significant genetic differentiation among *appalachiensis*, *glaucus* and

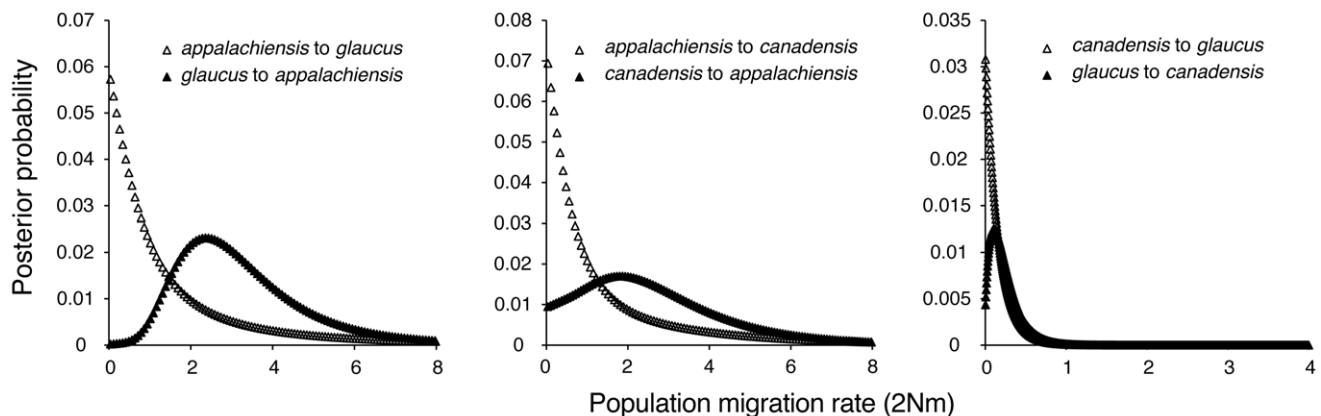


Figure 6. Estimated gene flow among *appalachiensis*, *glaucus*, and *canadensis*. Gene flow was estimated as the population migration rate or $2Nm$, which is equivalent to the historical average number of immigrants between species per generation: *glaucus* to *appalachiensis*: 2.3; *canadensis* to *appalachiensis*: 1.8; *appalachiensis* to either *glaucus* or *canadensis*: 0; *glaucus* to *canadensis*: 0.1; *canadensis* to *glaucus*: 0. doi:10.1371/journal.pgen.1002274.g006

Table 1. Estimates of linkage disequilibrium, arranged from highest to lowest values, with mean \pm SD.

AFLPs	Z-linked genes
laboratory-generated hybrids = 3.50 \pm 1.93	<i>canadensis</i> = 21.57 \pm 18.29 <i>appalachiensis</i> = 19.92 \pm 12.48
late-flight <i>canadensis</i> = 3.22 \pm 1.65	<i>glaucus</i> = 9.12 \pm 4.98
<i>canadensis</i> = 2.56 \pm 0.98	
<i>appalachiensis</i> = 2.35 \pm 1.02	
<i>glaucus</i> = 2.23 \pm 1.52	

Linkage disequilibrium was calculated as pair-wise associations among all polymorphisms, and summarized below as the average percentage of polymorphisms that were significantly ($p \leq 0.01$) associated with each other. For AFLPs, differences in levels of linkage disequilibrium among all the groups were highly significant ($p < 0.0001$), except among *appalachiensis* and *glaucus* ($p = 0.053$). For Z-linked genes, *canadensis* and *appalachiensis* were not significantly different from each other ($p = 0.214$) but they are both different from *glaucus* ($p < 0.0001$).

doi:10.1371/journal.pgen.1002274.t001

canadensis, and *appalachiensis* forms a very well-supported and distinct genomic cluster that is apparent in both the DNA sequence (Figure 2) and AFLP data (Figure 3, Table S5), (c) IMA2 simulations showed *appalachiensis* to have diverged from the parental species long ago (Figure 5), and (d) *Papilio appalachiensis* is parapatric with *canadensis* and occurs southward to Georgia in the southern Appalachian Mountains, very far outside *canadensis*' range (Figure 1A; Materials and Methods). Therefore, *appalachiensis* could not be a product of ongoing hybridization between *glaucus* and *canadensis*. It is also important to note that its phenotype is unlike that of the laboratory-generated *glaucus* x *canadensis* hybrids.

Conclusions

Our data show that *Papilio appalachiensis* displays multiple hallmarks of a hybrid species. Furthermore, it has two unique features that inform us on the dynamics of hybrid speciation. First, it may potentially be the sole example of a hybrid species that spatially overlaps with both parental species. Second, its sex-linked ecologically important traits, and therefore sex chromosome mosaicism, appear to be under selection. In this case the sharp

Table 2. Distribution of single nucleotide polymorphisms and haplotype diversity among *glaucus*, *canadensis*, and *appalachiensis*.

	Number of samples sequenced (number of sequences)	Number of single nucleotide polymorphisms in a species	Number (and %) of species-specific single nucleotide polymorphisms	Number of haplotypes in a species	Number (and %) of species-specific haplotypes
COI					
<i>canadensis</i>	24	11	7 (64%)	11	10 (91%)
<i>glaucus</i>	56	18	10 (56%)	19	15 (79%)
<i>appalachiensis</i>	28	8	3 (37%)	7	4 (57%)
KET					
<i>canadensis</i>	20 (22)	37	25 (68%)	21	17 (81%)
<i>glaucus</i>	18 (18)	20	18 (90%)	15	13 (87%)
<i>appalachiensis</i>	17 (19)	30	15 (50%)	11	8 (73%)
TH					
<i>canadensis</i>	21 (21)	6	4 (67%)	8	2 (25%)
<i>glaucus</i>	34 (34)	24	14 (58%)	27	25 (93%)
<i>appalachiensis</i>	35 (35)	23	8 (35%)	19	12 (63%)
TPI					
<i>canadensis</i>	25 (27)	57	53 (93%)	20	17 (85%)
<i>glaucus</i>	30 (30)	19	9 (47%)	22	20 (91%)
<i>appalachiensis</i>	16 (17)	21	9 (43%)	10	7 (70%)
PER					
<i>canadensis</i>	25 (25)	8	2 (25%)	10	4 (40%)
<i>glaucus</i>	32 (32)	5	3 (60%)	30	27 (90%)
<i>appalachiensis</i>	19 (19)	21	16 (76%)	9	5 (56%)
LDH					
<i>canadensis</i>	26 (26)	46	19 (41%)	26	23 (88%)
<i>glaucus</i>	28 (28)	38	23 (61%)	28	27 (96%)
<i>appalachiensis</i>	16 (16)	49	15 (31%)	14	11 (79%)
PAH					
<i>canadensis</i>	20 (27)	46	21 (46%)	26	23 (88%)
<i>glaucus</i>	20 (24)	109	43 (39%)	24	23 (96%)
<i>appalachiensis</i>	18 (19)	100	14 (14%)	15	12 (80%)

For some Z-linked genes, the number of sequences differed from the number of samples sequenced because of heterozygosity in males.

doi:10.1371/journal.pgen.1002274.t002

ecological clines along the latitudinal and altitudinal thermal gradients seem to maintain three, rather than two, spatially overlapping tiger swallowtail species. The evolution and persistence of *appalachiensis* in contact with its parental species suggests that hybridization among animals may result in selectively favored hybrid species that contribute to biodiversity. Moreover, sex chromosomes play an important role in speciation [43], and this has specifically been demonstrated in the Lepidoptera [44,45], but this is the first example in which sex-linked traits seem to have contributed to hybrid speciation.

As a whole, our data suggest a scenario for the origin and maintenance of *appalachiensis*. During one of the late Pleistocene glacial retreats, *canadensis* populations retreated from their southern range while *glaucus* populations advanced northward and upward into the mountains. The changing thermal landscape likely brought the advancing *glaucus* populations into contact with a relict *canadensis* population in the Appalachian Mountains. The ensuing hybridization seems to have been largely unidirectional, with *canadensis* males preferentially mating with *glaucus* females, as they do today [12]. Hence, *appalachiensis* is now fixed for the *glaucus* mitochondrial genome, along with its W-linked female-limited mimicry and dimorphism. Both these traits are under frequency-dependent selection in the present *appalachiensis* habitat because with glacial retreats, selection for mimicry has also moved northward along with the range of *B. philenor*. However, the cold Appalachian thermal habitat ensured that the Z-linked genes associated with this lifestyle persisted from its *canadensis* ancestry in the proto-*appalachiensis* populations [16,37]. This selection also persists today [38]. This unique combination of traits under varied selection may have been critical in helping *appalachiensis* evolve as a distinct species. The maintenance of genome-wide admixture in *appalachiensis* (Figure 3) also suggests that it was formed via a relatively brief burst of hybridization between the parental species, with little backcrossing with *glaucus* despite their continued sympatry throughout *appalachiensis*' range. A few other hybrid species have likely emerged in a similarly short span of time with potentially little backcrossing [8,46,47]. Interestingly, our estimated divergence times of approximately 100,000 and 90,000 years between *glaucus/canadensis* and *appalachiensis* (see Figure 5) fall precisely in the last interglacial period in North America, known as the Sangamonian Stage (125,000 to 75,000 years ago), which is congruent with the scenario just outlined. It may be possible to explore this scenario with additional population genetic data for these species.

Papilio appalachiensis may provide a rare genetic example of the creative role of hybridization in evolution [8]. Hybrids are often considered maladapted and viewed as lying in fitness valleys between the adaptive peaks that the parental species occupy in the adaptive landscape [3]. In case of tiger swallowtails, however, the northward and upward movement of the toxic *B. philenor* in the Appalachian Mountains may have created a new, unoccupied adaptive peak with selection for a combination of cold thermal habitat and mimicry. Neither of the parental species appears to be able to occupy this peak. Our results suggest an intriguing scenario in which the hybrid species (*appalachiensis*), with its precise combination of phenotypic traits, may have landed directly on this novel, unoccupied adaptive peak.

The primary goals of this work were to test whether *appalachiensis* exhibited genetic evidence of hybrid ancestry, mosaic sex chromosomes, and genome-wide admixture. Our results reveal several intriguing patterns that appear to support *appalachiensis*' status as a hybrid species and suggest an evolutionary scenario for its origin and maintenance. Our work also highlights specific areas where additional data will enrich our understanding of this system

in particular and the dynamics of hybrid speciation in general. One such area is identifying the specific mechanisms that generate reproductive isolation between *appalachiensis* and both of its parental species. Our genetic data, and the low frequency of *glaucus/appalachiensis* hybrids from areas where they co-occur in West Virginia, suggest that these two species exhibit some reproductive isolation. However, the ecological, behavioral and genetic factors that contribute to this isolation remain unknown, as does the exact degree of isolation. Specifically, we do not currently understand the mechanisms of isolation or the precise geographic ranges and areas of overlap between *appalachiensis* and *canadensis*. In addition, more widespread sampling of *appalachiensis* from across its range would allow us to test its genomic integrity throughout the range, and also to estimate the extent to which it hybridizes with each of the parental species. Given the latitudinal range of *appalachiensis* relative to *glaucus* and *canadensis*, it is possible that *appalachiensis* experiences variable amounts of recent introgression with each of the two parental species. Our current sampling was limited so we were unable to test this hypothesis. However, genetic data for our *appalachiensis* sample from its southern range in North Carolina and co-occurring samples of *appalachiensis* and *glaucus* from West Virginia suggest that *appalachiensis* may be fairly homogeneous throughout its range, but more widespread sampling is warranted. Scriber [48] has recently pointed out that ongoing climate warming may diminish the cool thermal mountain refuges in the southern Appalachian Mountains of northern Georgia, western North Carolina and eastern Tennessee in *appalachiensis*' southernmost range. The changing thermal landscape may induce increased introgression of *glaucus* genes into *appalachiensis*, which may diminish the genomic contribution from *canadensis* in these *appalachiensis* populations. Clearly, the rich biological detail of this system holds promise to test various aspects of hybrid speciation and persistence in animals.

Materials and Methods

Study System

The parental species, *P. glaucus* and *P. Canadensis*. Tiger swallowtails (*Papilio glaucus* species group) form a small, monophyletic species group of the American subgenus *Pterourus* of *Papilio* [40]. Many tiger swallowtail species look similar in general appearance, although they have consistent differences in ecological traits as well as larval, pupal and adult morphology [49,50]. Nonetheless, there had been much confusion about subspecies and species status of three parapatric taxa: *rutulus*, *canadensis* and *glaucus*, since they occasionally hybridize when in contact at their range margins [50] (also see Figure 1A). Due to the lack of complete sympatry and occasional hybridization, many earlier authors had treated *rutulus* and *canadensis* as subspecies of *glaucus* [25,50,51]. The three taxa have now been solidly established as distinct species through the past three decades of work by JMS and his collaborators. This work has revealed multiple mechanisms of reproductive isolation separating *glaucus* and *canadensis* that indicate that they are good biological species [12–25] (also see the Introduction).

Interspecific hybridization is common in nature and there are many well-studied interspecific hybrid zones [3,4,52,53]. Several features have been used to distinguish interspecific hybrid zones from contact zones between subspecies, most importantly: (a) interspecific hybrid zones are marked by a reduction in hybrids relative to the parental forms (“bimodal hybrid zone”) [54], and (b) the maintenance of “genotypic clusters” in spite of hybridization [55]. The *glaucus/canadensis* hybrid zone meets these two hallmarks of interspecific hybrid zones: (a) hybrids make up less

than 20% of the community in the very middle of the *glaucus/canadensis* hybrid zone [28], and (b) our present genetic data as well as previous allozyme work [29,56,57] show that *glaucus* and *canadensis* form distinct genotypic clusters despite hybridization. The rate of hybridization reported for *glaucus* and *canadensis* in the middle of the hybrid zone is within the range of rate of hybridization known for other well-established species pairs [52].

Lastly, more recent phylogenetic work has shown *rutulus* and *canadensis* to be well-separated from *glaucus* [40]. Our present work (Figure 2, Figure 3, Figure 4, Figure 5, Table S5) further reinforces the view that *rutulus*, *canadensis* and *glaucus* are genetically strongly diverged parapatric species, not subspecies, despite occasional hybridization.

The putative hybrid species, *P. appalachiensis*. *Papilio appalachiensis* has a distinct and readily recognizable phenotype spanning morphology, ecology and behavior, and this phenotype is maintained in sympatry/parapatry with the parental species, *glaucus* and *canadensis* [35,36]. There are several factors that suggest that *appalachiensis* is a distinct species rather than a recent hybrid population. For example, its melanic female form is unlike any other tiger swallowtail species, and both the male and female wing patterns are unlike any laboratory-generated *F1* and *F2* *glaucus* x *canadensis* hybrids. Similar to the *glaucus/canadensis* hybrid zone (see above), *appalachiensis* shows two hallmarks of a species: (a) it has a bimodal hybrid zone with *glaucus*, and (b) it forms a unique genetic cluster distinct from its parental species (see Figure 2 and Figure 3, and Results/Discussion). *Papilio appalachiensis* shows highly reduced rate of hybridization with *glaucus*, with which it completely overlaps in distribution (Figure 1A). Although *appalachiensis* is found at mid- and high elevations in the Appalachian Mountains and *glaucus* is usually found at lower elevations, our work shows that they co-occur widely both along the elevational and latitudinal gradients at least for several weeks during late May and June. For example, KK found fresh individuals of both species feeding side by side on the same flowering honeysuckle bushes below Spruce Knob, West Virginia, in May and June 2006. Thus, there is ample opportunity for hybridization between the two taxa. Nonetheless, we collected only 16 suspected *glaucus/appalachiensis* hybrids (phenotypically intermediate between *glaucus* and *appalachiensis*), along with 114 phenotypically pure *appalachiensis* and 172 phenotypically pure *glaucus* within the overlap zone of *appalachiensis* and *glaucus* from Tennessee to northeastern West Virginia. Thus, the proportion of suspected hybrids between the two species was very small, approximately 5.6%, making this a clearly bimodal hybrid zone. These numbers are well below the estimates of hybridization for many other hybridizing species [4] and point to reduced hybridization as one would expect for distinct species. As a clarification, individuals were scored as putative hybrids between *appalachiensis* and *glaucus* based on the diagnostic wing color patterns given in the original species description [35,36], mainly: (1) square versus crescent-shaped submarginal yellow spots on the hindwing, (2) contiguous yellow band versus row of spots on forewing underside, and (3) extent of blue on the upper hindwing of females. They were first identified as probable hybrids by KK and then all specimens were blindly scored by H. Pavulaan, who independently arrived at the same conclusion (Pavulaan originally described *appalachiensis*). If our species assignments had been wrong, underestimating the true rate of hybridization based on phenotypic characters, the genetic clustering shown in Figure 2 and Figure 3 would have been very different.

The contact zone and nature of isolation between *appalachiensis* and *canadensis* are currently not well understood. These two species are believed to be sympatric in the Appalachian Mountains from

West Virginia northward [35,58]. Almost all of our *appalachiensis* samples were collected near Spruce Knob, West Virginia (Table S1, Figure S5), very near localities where *canadensis* reportedly occurs. Based on our published [37,38] as well as unpublished work, we suspect that *canadensis* has recently moved northward along the Appalachian Mountains, perhaps completely out of West Virginia, due to ongoing climate warming. It is likely that *appalachiensis* and *canadensis* maintain a hybrid zone similar to the *glaucus-canadensis* hybrid zone, and this hybrid zone may also have moved northward due to the ongoing warming in this area. Current work by MLA and JMS aims to delineate the present northern boundary of *appalachiensis* distribution, the southern boundary of *canadensis* in the Appalachian Mountains, and the boundaries of the contact zone between the two species. However, based on current understanding [35,58], *appalachiensis* and *canadensis* are narrowly sympatric or parapatric.

Specimen Collection and Lab Methods

Table S1 lists details of the 244 wild-caught individuals and 23 laboratory-generated hybrids used for genetic analysis, and Figure S5 shows localities where these specimens were collected. We preserved bodies in ethanol and stored wings in glassine envelopes. We measured wingspan from the preserved wings using vernier calipers, from the base of the forewing to its tip (Figure S1). We extracted genomic DNA from legs and thoracic muscle tissue using DNeasy Blood & Tissue Kit (Qiagen, Germantown, MD).

We sequenced one mitochondrial (*Cytochrome oxidase I, COI*) and six Z-linked genes (*Kettin (Ket)*, *Tyrosine hydroxylase (TH)*, *Triosephosphate isomerase (Tpi)*, *Period (Per)*, *Lactate dehydrogenase (Ldh)* and *Phenylalanine hydroxylase (PAH)*) to study potential admixture in *appalachiensis* (NCBI accession numbers JF764373-JF764558, JF951433-JF951722). Primer sequences and lengths of aligned sequences are given in Table S2. PCR protocols used for previously published primers are given in the literature [27,59]. We designed primers for *TH* and *PAH* by aligning sequences from *Bombyx mori* and *Papilio xuthus* (Table S2). Previous allozyme work had shown that *appalachiensis* has a *glaucus*-like *Pgd* allele [38], so we also developed several primers for that gene. Unfortunately, none of these primers worked. Future comparisons of *Pgd* sequences among the focal species will be useful for inferring the extent of mosaicism along the *appalachiensis* Z chromosome. We amplified all Z-linked genes with Invitrogen Platinum hot-start taq (Invitrogen Corp, Carlsbad, CA) and a touchdown protocol with annealing temperatures from 65–50°C.

We generated amplified fragment length polymorphism (AFLP) data using a Plant Mapping Kit (Applied Biosystems, Foster City, CA) and 10 primer pairs: EcoRI-ACA/MseI-CAC, EcoRI-ACA/MseI-CAG, EcoRI-ACA/MseI-CAT, EcoRI-ACA/MseI-CTC, EcoRI-ACA/MseI-CTG, EcoRI-ACT/MseI-CAC, EcoRI-ACT/MseI-CAG, EcoRI-ACT/MseI-CAT, EcoRI-ACT/MseI-CTC, and EcoRI-ACT/MseI-CTG.

Analyses of DNA Sequence Data

We cleaned DNA sequences using Sequencher (Gene Codes Corp., Ann Arbor, MI). When available, we included reference sequences from GenBank and aligned the sequence datasets online using Multiple Sequence Comparison by Log-Expectation program (MUSCLE; <http://www.ebi.ac.uk/Tools/muscle/index.html>). We corrected the sequence output alignments by eye, paying special attention to indels in introns of Z-linked genes. For population genetic analyses, indels longer than one base pair were scored as a single polymorphism.

We used this dataset to analyze sequence similarity between *glaucus*, *canadensis* and *appalachiensis* by performing species pair-wise

locus-by-locus analyses of molecular variance (AMOVAs) using Arlequin [60] to calculate fixation index (F_{ST}) values (Figure 2B, Table S4). Given their young ages and occasional hybridization, there is considerable polymorphism but few fixed differences among the three species. Hence, we tested the genetic ancestry of *appalachiensis* in relation to the parental species for polymorphisms that were strongly differentiated ($p \leq 0.001$) between *glaucus* and *canadensis* (Figure 2A). However, we used entire sequences to calculate the F_{ST} values in Figure 2B. Specimens that were sequenced for only one gene are included in Table S3 (missing data are marked grey) and were used in calculating F_{ST} values in Figure 2B, but excluded from Figure 2A. Genotypes here refer to nucleotide bases at specific SNP or indel polymorphisms in the genes, which can be diploid (Z-linked polymorphisms scored in males) or haploid (mtDNA, and Z-linked polymorphisms scored in females). The Z-linked genes in Figure 2A and Tables S2, S3, S4 are listed in the same order in which they occur on the *Bombyx mori* Z chromosome.

The Use of Mitochondrial DNA to Infer the History of the W Chromosome

We wanted to determine whether *appalachiensis* inherited its W chromosome, where the gene controlling female-limited mimicry is situated, from *canadensis* or *glaucus*. Unfortunately, no W-linked molecular markers exist for these butterflies. Therefore, we used a mitochondrial gene because the W chromosome and the mitochondrial genome are both maternally inherited and thus linked in Lepidoptera. Andolfatto *et al.* [61] had attempted to use mtDNA to examine the evolutionary history of the W-linked female-limited mimetic phenotype in *glaucus*. Surprisingly, they found no association between mtDNA genealogies and mimicry, and attributed this finding to paternal leakage of mtDNA and recombination among maternal and paternal mtDNA. In fact, there are three other, potentially more likely explanations for the polyphyly of mimicry with respect to the mitochondrial gene tree: (1) multiple origins of mimicry resulting in more than one W-linked mimicry allele, (2) loss of mimicry in some individuals resulting in more than one non-mimicry allele, and (3) the presence of a low-frequency Z-linked mimicry suppressor allele in *glaucus* [26,61]. Thus, the observed disassociation between mtDNA and mimicry does not necessarily mean disassociation between mtDNA and the W chromosome. Unfortunately, the above alternatives cannot be rigorously tested at present because the molecular identities of the mimicry and enabler genes are yet unknown. However, it is important to remember that there is overwhelming evidence from throughout the animal kingdom showing that mtDNA is almost exclusively maternally inherited [62–64]. Finally, it is critical to note that rare paternal inheritance of mtDNA could only influence our interpretation regarding the history of the W chromosome if it occurred during the specific time period during which *glaucus* and *canadensis* hybridized to form *appalachiensis*. If this had occurred, we would predict a mixture of *glaucus* and *canadensis* mtDNA haplotypes in *appalachiensis*, which we do not see (Figure 2A). Instead, our extensive dataset reveals that *glaucus* and *canadensis* mtDNA haplotypes are exclusively species-specific, and that of *appalachiensis* is like *glaucus* (Figure 2A). This pattern shows that paternal leakage of mtDNA has not occurred between *glaucus* and *canadensis* during the origin of *appalachiensis*. Thus, mtDNA is an informative marker for inferring the species-level ancestry of *appalachiensis*' W chromosome.

Analyses of AFLP Data

We analyzed AFLP data using Genemapper Software Version 4.0 (Applied Biosystems, Foster City, CA). Samples were analyzed

with the default bin width of 1.0 bp, and the fragment analysis range of 50–250 bps. We analyzed the AFLP data in two ways: we identified population clusters and genomic admixture among the focal species using the program STRUCTURE [65], and we constructed a phylogeny of tiger swallowtails using PAUP* [66]. We performed population clustering in STRUCTURE using three datasets: (a) focal species alone (*glaucus* (39 samples), *canadensis* (41 samples) and *appalachiensis* (38 samples); total 118 samples) (Figure S2A), (b) focal species along with laboratory-generated *glaucus/canadensis* hybrids (23 hybrid samples; total 141 samples) (Figure 3A), and (c) focal species along with laboratory-generated *glaucus/canadensis* hybrids and wild caught late flight *canadensis* individuals (30 late flight *canadensis* individuals; total 171 samples) (Figure 3C, Figure S2B).

We performed three separate STRUCTURE analyses because each was aimed at testing a different hypothesis and thus required that a different subset of samples be included. For instance, our first analysis focused only on the three focal species to determine whether *appalachiensis* exhibited evidence of admixture between *canadensis* and *glaucus* yet was genetically distinct from both of them (Figure S2A). To do this, we used a well-established approach that has been used previously to both support and refute the hypothesis of hybrid speciation [6,67]. When admixture is present, as in the case of hybrid speciation, STRUCTURE can estimate the proportion of each individual's genome that is derived from distinct gene pools (admixture proportions). Gompert *et al.* [6] showed that a hybrid species with genomic mosaicism displayed admixture proportions split between the parental species when analyzed with the parental species at $K = 2$, but then formed a distinct cluster at $K = 3$ (just as we have shown here for *appalachiensis*) [6]. In comparison, Kronforst *et al.* [67] showed that if one does the same analysis with a species that does not appear to be a hybrid species, at $K = 2$ it clusters entirely with its sister species but then forms its own cluster at $K = 3$ [67]. Thus, we used STRUCTURE to estimate admixture proportions instead of assignment probabilities of the sampled individuals. We did this because we were asking what proportion of genome of each *appalachiensis* individual was *glaucus*- versus *canadensis*-like (which is shown by admixture proportions), not whether a particular *appalachiensis* individual was assigned to one or the other parental species with a certain probability (which is shown by assignment probabilities). Using this method, we discovered a pattern in *appalachiensis* (Figure 3) consistent with it being a hybrid species [6]. We used a similar approach for our second analysis focused on comparing the genetic signature of *appalachiensis* to lab generated *glaucus/canadensis* hybrids (Figure 3A), while our third analysis focused on the ancestry of the late-flight *canadensis* population (Figure 3C).

Our STRUCTURE analyses were based on 249 AFLP markers that were significantly differentiated ($p \leq 0.05$) between *glaucus* and *canadensis* as judged by F_{ST} values from a locus-by-locus AMOVA. We analyzed each dataset in STRUCTURE with these ancestry-informative AFLP markers with 200,000 burn-in generations and 1,000,000 generations of data collection under the assumption of two, three and four admixed populations ($K = 2, 3$ or 4). We used STRUCTURE to estimate admixture proportions rather than assignment probabilities, and visualized the STRUCTURE outputs with the program Distruct [68] (Figure 3). Analyses based on the full AFLP dataset (polymorphisms with a minor allele frequency $\geq 5\%$) produced similar results but with less resolution (Figure S3).

For Figure 3B, we used the same 249 AFLP markers that were significantly differentiated between *glaucus* and *canadensis* (see last paragraph), did pair-wise AMOVAs among the three species, and

categorized AFLP loci as: 1) “*glaucus*-like”: *appalachiensis* allele frequency is significantly different from *canadensis* but not from *glaucus*, 2) “*canadensis*-like”: *appalachiensis* allele frequency is significantly different from *glaucus* but not from *canadensis*, 3) “intermediate”: *appalachiensis* allele frequency is intermediate but not significantly different from either parental species, and 4) “different”: *appalachiensis* allele frequency is different from both parental species. The pie chart (Figure 3B) shows the number of *appalachiensis* AFLP loci that fall in the four categories, highlighting the intermediacy of *appalachiensis* with respect to *glaucus* and *canadensis*.

Phylogenetic Methods

In order to study the larger evolutionary context of *appalachiensis*' unique hybrid phenotype, we used the AFLP dataset to resolve phylogenetic relationships between tiger swallowtails. The AFLP phylogenetic dataset included 184 individuals with *P. garamas* as the outgroup, and 1,607 polymorphic AFLP markers. Laboratory-generated *glaucus/canadensis* hybrids and suspected wild-caught *glaucus/appalachiensis* hybrids were excluded from the phylogenetic analysis. We generated a neighbor-joining tree of tiger swallowtails in PAUP* [66] (Figure 4A), with bootstrap support based on 2,000 pseudo-replicates. The distance-based neighbor-joining method was used because there is no model of evolution for AFLP data [69–73]. AFLPs are reliable for phylogenetic reconstruction and, in fact, provide answers consistent with other sources of data in addition to being able to resolve relationships that DNA sequence data cannot [74–76]. Zakharov *et al.* [40] only included *glaucus*, *canadensis*, *rutulus* and *multicaudata*, so only half of the eight tiger swallowtail species were covered in their phylogeny. Note that our relationships among those species are exactly the same as in Zakharov *et al.*, which shows that AFLPs resolve phylogenies consistent with other sources of molecular data. The critical species for inferring the evolution of mimicry among tiger swallowtails are *alexiares* and *garcia*, which we included in our phylogeny for the first time. Several nuclear and mitochondrial genes tested were not informative in resolving relationships between some tiger swallowtail species (analysis not presented here), so fast-evolving molecular markers such as AFLPs were needed to resolve these relationships.

We mapped on this phylogeny the ecological and morphological traits (Table S6) relevant to the hybrid phenotype of *appalachiensis* using the software MacClade 4 [77]. Since phylogenetic programs do not model speciation and trait evolution on reticulated trees, for the initial mapping of traits we excluded *appalachiensis* from the tree and evolved traits from the outgroup (*P. garamas*) to *glaucus* and *canadensis*. Based on the signatures of hybrid speciation in Figure 2 and Figure 3, *appalachiensis* and its mosaic inheritance of traits from *glaucus* and *canadensis* were subsequently manually added to the tree, post-hoc.

Estimating Divergence Times and Gene Flow

We used our DNA sequence data and the program IMA2 [78–81] to estimate divergence times (Figure 5) and historical gene flow (Figure 6) among *appalachiensis*, *glaucus* and *canadensis*. This coalescent-based method uses comparative DNA sequence data and an “Isolation with Migration” model of population divergence to estimate model parameters such as effective population sizes, divergence times, and bi-directional migration rates. While IMA2 can analyze more than two populations at a time, it requires populations to be related by a specified bifurcating tree. Since we were testing the predicted hybrid origin of *appalachiensis*, which does not follow a bifurcating mode of speciation, we analyzed our data as three pair-wise comparisons among the three focal species. For each analysis, we ran IMA2 with 100 Metropolis-coupled

chains with a 150,000-step burn-in followed by 20 million steps of data collection. An important requirement of the IMA2 method is that there is free recombination among sequenced loci: loci do not necessarily have to be on separate chromosomes but they should not be too close to one another. Since six of our seven sequenced loci were on the Z chromosome, and they spanned the length of that chromosome, we satisfied this requirement by using data from only three genes for these analyses: *Kettin* and *PAH*, which are on opposite ends of the Z chromosome, and the mitochondrial gene *COI*. Another requirement of IMA2 is that there is no recombination within loci. We satisfied this requirement by trimming the sequence dataset to a region with no evidence of recombination using the program IMgc Online [82]. We used this trimmed dataset with three genes for estimating both divergence times and gene flow. Converting parameter estimates to units such as time in years since divergence or population migration rates ($2Nm$) requires mutation rates for each gene. We estimated these by comparing sequences to the within tiger swallowtail outgroup, *P. multicaudata*. Average mtDNA divergence between *multicaudata* and the three focal species was 3.17%. Using previous *COI* divergence estimate of 1.15% per lineage per million years [83], we estimated that *P. multicaudata* split from the common ancestor of our focal species approximately 1.4 million years ago. This date differs from that estimated by Zakharov *et al.* [40] for all of *Papilio*, but our date may be more accurate since it was estimated specifically for tiger swallowtails. However, the actual dates are less important for our purpose since the relevant findings from the IMA2 analysis were that: (a) *appalachiensis* originated much more recently compared to the time when *glaucus* and *canadensis* split from one another (Figure 5), and (b) *appalachiensis* split from both *glaucus* and *canadensis* around the same time (Figure 5).

A similar analysis was done to estimate divergence times between *glaucus*, *canadensis* and late-flight *canadensis* (Figure S4).

Comparisons of Linkage Disequilibrium

We estimated the extent of linkage disequilibrium across both our AFLP and Z-linked DNA sequence data and compared this among populations. For each dataset (AFLPs or concatenated Z-linked gene sequences) we calculated pair-wise associations among all polymorphisms using the exact test for linkage disequilibrium implemented in Arlequin. We summarized these results by calculating the average percentage of polymorphisms that were significantly ($p \leq 0.01$) associated with each other (Table 1) and then compared them among species using an ANOVA.

Comparisons of Species-Specific Polymorphisms

We used our DNA sequence dataset to calculate the number of species-specific polymorphisms for each of our focal species (Table 2). We calculated species-specific polymorphisms as the proportion of genetic polymorphisms unique to one species, out of the total number of polymorphisms (including shared and species-specific) seen among all the sequenced samples of that species. We used the software DnaSP [84] to count the total number of polymorphisms and the subset that were species-specific for each species at each gene. We then compared the average proportion of species-specific polymorphisms among species with an ANOVA. We also performed a similar analysis comparing the proportion of species-specific haplotypes rather than polymorphisms, and found similar results (Table 2).

Supporting Information

Figure S1 Distribution of wingspan (representing body size) among *glaucus*, *appalachiensis* and *canadensis* (mean \pm SD): *glaucus*:

57±6.11 mm, n = 45; *appalachiensis*: 57±3.11 mm, n = 45; *canadensis*: 47±2.04 mm, n = 23. ANOVA: $F_{(2,110)} = 45.588$; $p < 0.0001$.

(TIF)

Figure S2 STRUCTURE analysis comparing (A) *appalachiensis* with the parental species, and (B) *appalachiensis* with lab hybrids and late flight *canadensis*.

(TIF)

Figure S3 STRUCTURE analysis with all polymorphic AFLP markers with a minor allele frequency > 5%.

(TIF)

Figure S4 Estimated divergence times between *glaucus*, *canadensis* and late flight *canadensis*. Dates of divergence estimated by IMA2 are: (a) late flight *canadensis* and *glaucus*: approximately 400,000 years ago, (b) late flight *canadensis* and *canadensis*: approximately zero years ago.

(TIF)

Figure S5 Localities and states where specimens used in this study were collected. For each locality, the number of specimens collected of each focal species is shown. can = *Papilio canadensis*, late flight can = *canadensis* late flight, app = *P. appalachiensis*, gla = *P. glaucus*, and gla-app suspected hybrids = individuals phenotypically intermediate between *glaucus* and *appalachiensis*, hence suspected to be hybrids between the two species.

(TIF)

Table S1 Specimens used in this work.

(PDF)

Table S2 Genes, primer sequences, and lengths of aligned sequences used in this work.

(DOC)

Table S3 Complete genotype table for Figure 2.

(PDF)

References

- Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, et al. (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301: 1211–1216.
- Lewontin RC, Birch LC (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution* 20: 315–336.
- Mallet J (2007) Hybrid speciation. *Nature* 446: 279–283.
- Mallet J (2005) Hybridization as an invasion of the genome. *Trends Ecol Evol* 20: 229–237.
- Coyne JA, Orr HA (2004) Speciation. Sunderland Massachusetts: Sinauer Associates.
- Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC (2006) Homoploid hybrid speciation in an extreme habitat. *Science* 314: 1923–1925.
- Mavarez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, et al. (2006) Speciation by hybridization in *Heliconius* butterflies. *Nature* 441: 868–871.
- Nolte AW, Tautz D (2010) Understanding the onset of hybrid speciation. *Trends Gen* 26: 54–58.
- Gross BL, Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. *J Heredity* 96: 241–252.
- Salazar C, Baxter SW, Pardo-Diaz C, Wu G, SurrIDGE A, et al. (2010) Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. *PLoS Genet* 6: e1000930. doi:10.1371/journal.pgen.1000930.
- Mavarez J, Linares M (2008) Homoploid hybrid speciation in animals. *Mol Ecol* 17: 4181–4185.
- Deering MD, Scriber JM (2002) Field bioassays show heterospecific mating preference asymmetry between hybridizing North American Papilio butterfly species (Lepidoptera: Papilionidae). *J Ethol* 20: 25–33.
- Scriber JM, Deering M, Stump A (2003) Hybrid zone ecology and tiger swallowtail trait clines in North America. In: Boggs CL, Watt WB, Ehrlich PR, eds. *Butterflies: Ecology and Evolution Taking Flight*. Chicago: University of Chicago Press. pp 367–391.
- Scriber JM, Lederhouse RC, Dowell RV (1995) Hybridization studies with North American swallowtails. In: Scriber JM, Tsubaki Y, Lederhouse RC, eds. *Swallowtail Butterflies: Their Ecology and Evolutionary Biology*. Gainesville, Florida: Scientific Publishers, Inc. pp 269–281.
- Hagen RH, Scriber JM (1995) Sex chromosomes and speciation in tiger swallowtails. In: Scriber JM, Tsubaki Y, Lederhouse RC, eds. *Swallowtail Butterflies: Their Ecology and Evolutionary Biology*. Gainesville, Florida: Scientific Publishers, Inc. pp 211–227.
- Mercader RJ, Scriber JM (2008) Asymmetrical thermal constraints on the parapatric species boundaries of two widespread generalist butterflies. *Ecol Entomol* 33: 537–545.
- Lederhouse RC, Ayres MP, Scriber JM (1995) Physiological and behavioral adaptations to variable thermal environments in North American swallowtail butterflies. In: Scriber JM, Tsubaki Y, Lederhouse RC, eds. *Swallowtail Butterflies: Their Ecology and Evolutionary Biology*. Gainesville, Florida: Scientific Publishers, Inc. pp 71–81.
- Scriber JM, Lederhouse RC (1992) The thermal environment as a resource dictating geographic patterns of feeding specialization of insect herbivores. In: Hunter MR, Ohgushi T, Price PW, eds. *Effects of Resource Distribution on Animal-Plant Interactions*. San Diego California: Academic Press. pp 429–466.
- Lindroth RL, Scriber JM, Hsia MTS (1986) Differential responses of tiger swallowtail subspecies to secondary metabolites from tulip tree and quaking aspen. *Oecologia* 70: 13–19.
- Mercader RJ, Aardema ML, Scriber JM (2009) Hybridization leads to host-use divergence in a polyphagous butterfly sibling species pair. *Oecologia* 158: 651–662.
- Mercader RJ, Scriber JM (2007) Diversification of host use in two polyphagous butterflies: differences in oviposition specificity or host rank hierarchy. *Ent Exp Appl* 125: 89–101.
- Bossart JL, Scriber JM (1995) Genetic variation in oviposition preference in tiger swallowtail butterflies: interspecific, interpopulation and interindividual comparisons. In: Scriber JM, Tsubaki Y, Lederhouse RC, eds. *Swallowtail Butterflies: Their Ecology and Evolutionary Biology*. Gainesville, Florida: Scientific Publishers, Inc. pp 183–193.

Table S4 A comparison of locus-by-locus AMOVA among genes and species. F_{ST} values and their statistical significance are shown for each gene and each species pair-wise comparison. Also shown are the numbers of loci (i.e., genetic polymorphisms) that were significantly different either at $p \leq 0.05$ or at $p \leq 0.001$. This table has been condensed into Figure 2B.

(DOC)

Table S5 Pair-wise population differentiation among tiger swallowtails based on AFLP data. Overall F_{ST} : 0.16, $p < 0.001$. Population pair-wise F_{ST} .

(DOC)

Table S6 Ecological and morphological differentiation between tiger swallowtails.

(DOC)

Acknowledgments

Thanks to Harry Pavulaan and David Wright for advice; Nicola Chamberlain, Thomas Dakermanji, Swee Peck Quek, and Jean Santos for assistance in lab work and phylogenetic methods; and Peter Andolfatto, Charles Marshall, Ulrich Mueller, and Felix Sperling for comments on the manuscript. The following colleagues contributed specimens to this work, assisted in specimen collection, or provided collecting permits: Ken Davenport, Alex Grkovich, Dan Hardy, Rick Patterson, Harry Pavulaan, H. Romack, Andrew Warren, David M. Wright (USA), Sergio Sanchez-Pena, Patricia Soriano and the staff at the Parque Ecologica Chipinque (Monterrey, Mexico), Armando Luis-Martínez, José Asael Nájera Carpio, Nancy Araceli Alvarado Albino, Alonso Turrent, Rafael Turrent (Mexico City), Jesus Garcia and Irmi Garcia (Ciudad Victoria, Mexico). Andrew Warren provided photographs of *P. alexiares*, which are copyright of the De La Maza Family Collection, Mexico City.

Author Contributions

Conceived and designed the experiments: KK TEJ LEG MRK. Performed the experiments: KK CS MRK. Analyzed the data: KK CS MRK. Contributed reagents/materials/analysis tools: MLA JMS. Wrote the paper: KK MRK.

23. Ritland DB, Scriber JM (1985) Larval developmental rates of three subspecies of tiger swallowtail butterflies, *Papilio glaucus*, and their hybrids in relation to temperature. *Oecologia* 65: 185–193.
24. Scriber JM, Lindroth RL, Nitao JK (1989) Differential toxicity of a phenolic glycoside from quaking aspen to *Papilio glaucus* butterfly subspecies, hybrids and backcrosses. *Oecologia* 81: 186–191.
25. Hagen RH, Lederhouse RC, Bossart JL, Scriber JM (1991) *Papilio canadensis* and *P. glaucus* (Papilionidae) are distinct species. *J Lep Soc* 45: 245–258.
26. Scriber JM, Hagen RH, Lederhouse RC (1996) Genetics of mimicry in the tiger swallowtail butterflies, *Papilio glaucus* and *P. canadensis* (Lepidoptera: Papilionidae). *Evolution* 50: 222–236.
27. Putnam AS, Scriber JM, Andolfatto P (2007) Discordant divergence times among Z-chromosome regions between two ecologically distinct swallowtail butterfly species. *Evolution* 61: 912–927.
28. Luebke HJ, Scriber JM, Yandell BS (1988) Use of Multivariate Discriminant Analysis of male wing morphometrics to delineate a hybrid zone for *Papilio glaucus glaucus* and *P. g. canadensis* in Wisconsin. *Amer Midland Natur* 119: 366–379.
29. Hagen RH, Scriber JM (1989) Sex-linked diapause, color, and allozyme loci in *Papilio glaucus*: linkage analysis and significance in a hybrid zone. *J Heredity* 80: 179–185.
30. Ries L, Mullen SP (2008) A rare model limits the distribution of its more common mimic: A twist on frequency-dependent Batesian mimicry. *Evolution* 62: 1798–1803.
31. Kunte K (2009) Female-limited mimetic polymorphism: A review of theories and a critique of sexual selection as balancing selection. *Anim Behav* 78: 1029–1036.
32. Brower JVZ (1958) Experimental studies of mimicry in some North American butterflies: Part II. *Battus philenor* and *Papilio troilus*, *P. polyxenes* and *P. glaucus*. *Evolution* 12: 123–136.
33. Brower LP, Brower JVZ (1962) The relative abundance of model and mimic butterflies in natural populations of the *Battus philenor* mimicry complex. *Ecology* 43: 154–158.
34. Burns JM (1966) Preferential mating versus mimicry: Disruptive selection and sex-limited dimorphism in *Papilio glaucus*. *Science* 153: 551–553.
35. Pavulaan H, Wright DM (2002) *Pterourus appalachiensis* (Papilionidae: Papilioninae), a new swallowtail butterfly from the Appalachian Region of the United States. *Tax Rep Internat Lep Sur* 3: 1–20.
36. Pavulaan H, Wright DM (2004) Discovery of a black female form of *Pterourus appalachiensis* (Papilionidae: Papilioninae) and additional observations of the species in West Virginia. *Tax Rep Internat Lep Sur* 6: 1–10.
37. Scriber JM, Ordng GJ (2005) Ecological speciation without host plant specialization; possible origins of a recently described cryptic *Papilio* species. *Ent Exp Appl* 115: 247–263.
38. Scriber JM, Ordng GJ, Mercader RJ (2008) Introgression and parapatric speciation in a hybrid zone. In: Tilmon KJ, ed. *Specialization, Speciation, and Radiation: The Evolutionary Biology of Herbivorous Insects*. Berkeley, California: University of California Press. pp 69–87.
39. Sperling FAH (2003) Butterfly molecular systematics: From species definitions to higher-level phylogenies. In: Boggs CL, Watt WB, Ehrlich PR, eds. *Butterflies: Ecology and Evolution Taking Flight*. Chicago: University of Chicago Press. pp 431–458.
40. Zakharov EV, Caterino MS, Sperling FAH (2004) Molecular phylogeny, historical biogeography, and divergence time estimates for Swallowtail Butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Syst Biol* 53: 193–215.
41. Ordng GJ, Mercader RJ, Aardema ML, Scriber JM (2010) Allochronic isolation and incipient hybrid speciation in tiger swallowtail butterflies. *Oecologia* 162: 523–531.
42. Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: Harrison RG, ed. *Hybrid Zones and the Evolutionary Process*. USA: Oxford University Press. pp 13–45.
43. Qvarnström A, Bailey RI (2009) Speciation through evolution of sex-linked genes. *Heredity* 102: 4–15.
44. Sperling FAH (1994) Sex-linked genes and species differences in Lepidoptera. *Can Ent* 126: 807–818.
45. Prowell DP (1998) Sex linkage and speciation in Lepidoptera. In: Howard DJ, Berlocher SH, eds. *Endless forms: Species and speciation*. Oxford: Oxford University Press. pp 309–319.
46. Schwarz D, Matta BM, Shakir-Botteri NL, McPherson BA (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* 436: 546–549.
47. Nolte AW, Freyhof J, Stenshorn KC, Tautz D (2005) An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proc R Soc B* 272: 2379–2387.
48. Scriber JM (2011) Impacts of climate warming on hybrid zone movement: Geographically diffuse and biologically porous “species borders”. *Insect Sci* 18: 121–159.
49. Brower LP (1959) Speciation in butterflies of the *Papilio glaucus* group. I. Morphological relationships and hybridization. *Evolution* 13: 40–63.
50. Scott JA (1986) *The Butterflies of North America: A Natural History and Field Guide*. Stanford, USA: Stanford University Press.
51. Rothschild W, Jordan K (1906) A revision of the American *Papilios*. *Novitates Zoologicae* 13: 411–744.
52. Mallet J, Beltran M, Neukirchen W, Linares M (2007) Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evol Biol* 7: 28.
53. Sperling FAH (1990) Interspecific hybrids in *Papilio* butterflies: Poor taxonomy or interesting evolutionary problem? *Can J Zool* 68: 1790–1799.
54. Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends Ecol Evol* 15: 250–255.
55. Mallet J (1995) A species definition for the modern synthesis. *Trends Ecol Evol* 10: 294–299.
56. Hagen RH, Scriber JM (1991) Systematics of the *Papilio glaucus* and *P. troilus* species groups (Lepidoptera: Papilionidae): inferences from allozymes. *Annals Ent Soc Amer* 84: 380–395.
57. Hagen RH (1990) Population structure and host use in hybridizing subspecies of *Papilio glaucus* (Lepidoptera: Papilionidae). *Evolution* 44: 1914–1930.
58. Allen TJ (1997) *The butterflies of West Virginia and their caterpillars* University of Pittsburgh Press.
59. Nazari V, Zakharov EV, Sperling FAH (2007) Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology and seven genes. *Mol Phylogenet Evol* 42: 131–156.
60. Excoffier L, Lischer H (2009) Arlequin: An Integrated Software Package for Population Genetics Data Analysis. Ver 3.11.
61. Andolfatto P, Scriber JM, Charlesworth B (2003) No association between mitochondrial DNA haplotypes and a female-limited mimicry phenotype in *Papilio glaucus*. *Evolution* 57: 305–316.
62. Birky CW, Jr. (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci, USA* 92: 11331–11338.
63. Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics* 18: 269–292.
64. Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, et al. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann Rev Ecol Syst* 18: 489–522.
65. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
66. Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, MA: Sinauer Associates.
67. Kronforst MR, Salazar C, Linares M, Gilbert LE (2007) No genomic mosaicism in a putative hybrid butterfly species. *Proc R Soc B* 274: 1255–1264.
68. Rosenberg NA (2007) Distruct: A Program for the Graphical Display of Population Structure. Ver 1.1.
69. Albertson RC, Markert JA, Danley PD, Kocher TD (1999) Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. *Proc Natl Acad Sci, USA* 96: 5107–5110.
70. Beardsley PM, Yen A, Olmstead RG (2003) AFLP phylogeny of *Mimulus* section *Erythranthe* and the evolution of hummingbird pollination. *Evolution* 57: 1397–1410.
71. Hollingsworth PM, Ennos RA (2004) Neighbour joining trees, dominant markers and population genetic structure. *Heredity* 92: 490–498.
72. Mendelson TC, Shaw KL (2005) Rapid speciation in an arthropod. *Nature* 433: 375–376.
73. Rudh A, Rogell B, Hoglund J (2007) Non-gradual variation in colour morphs of the strawberry poison frog *Dendrobates pumilio*: genetic and geographical isolation suggest a role for selection in maintaining polymorphism. *Mol Ecol* 16: 4284–4294.
74. Chamberlain NL, Hill RI, Kapan DD, Gilbert LE, Kronforst MR (2009) Polymorphic butterfly reveals the missing link in ecological speciation. *Science* 326: 847–850.
75. Quek S-P, Counterman BA, Moura PAD, Cardoso MZ, Marshall CR, et al. (2010) Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. *Proc Natl Acad Sci, USA* 107: 7365–7370.
76. Savage WK, Mullen SP (2009) A single origin of Batesian mimicry among hybridizing populations of admiral butterflies (*Limenitis arthemis*) rejects an evolutionary reversion to the ancestral phenotype. *Proc R Soc B* 276: 2557–2565.
77. Maddison DR, Maddison WP (2003) MacClade 4. Version 4.06. Sunderland, MA: Sinauer Associates.
78. Nielsen R, Wakeley J (2001) Distinguishing migration from isolation. A Markov chain Monte Carlo approach. *Genetics* 158: 885–96.
79. Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.
80. Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Mol Biol Evol* 22: 297–307.
81. Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov Chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci, USA* 104: 2785–2790.
82. Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics* 23: 1851–1853.
83. Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci, USA* 91: 6491–6495.
84. Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.