Repeated Reticulate Evolution in North American *Papilio machaon* Group Swallowtail Butterflies

Julian R. Dupuis*, **Felix A. H. Sperling**

Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

* jrdupuis@ualberta.ca

Abstract

Hybridization between distinct populations or species is increasingly recognized as an important process for generating biodiversity. However, the interaction between hybridization and speciation is complex, and the diverse evolutionary outcomes of hybridization are difficult to differentiate. Here we characterize potential hybridization in a species group of swallowtail butterflies using microsatellites, DNA sequences, and morphology, and assess whether adaptive introgression or homoploid hybrid speciation was the primary process leading to each putative hybrid lineage. Four geographically separated hybrid populations were identified in the *Papilio machaon* species group. One distinct mitochondrial DNA clade from *P. machaon* was fixed in three hybrid taxa (*P. brevicauda*, *P. joanae*, and *P. m. kahli*), while one hybrid swarm (*P. zelicaon* x *machaon*) exhibited this hybrid mtDNA clade as well as widespread parental mtDNA haplotypes from both parental species. Microsatellite markers and morphology showed variable admixture and intermediacy, ranging from signatures of prolonged differential introgression from the paternal species (*P. polyxenes*/*P. zelicaon*) to current gene flow with both parental species. Divergences of the hybrid lineages dated to early- to mid-Pleistocene, suggesting that repeated glaciations and subsequent range shifts of parental species, particularly *P. machaon hudsonianus*, facilitated initial hybridization. Although each lineage is distinct, *P. joanae* is the only taxon with sufficient evidence (ecological separation from parental species) to define it as a homoploid hybrid species. The repetition of hybridization in this group provides a valuable foundation for future research on hybridization, and these results emphasize the potential for hybridization to drive speciation in diverse ways.

Introduction

Hybridization between distinct populations or species has historically been considered trivial in the overall diversification of animal life (e.g. [1]) or as a countervailing force to speciation [2]. Modern molecular methods have turned this notion on its head, and it is becoming increasingly clear that hybridization is both more prevalent and evolutionarily more important

than previously thought [3, 4, 5, 6, 2]. Hybridization can have a continuum of complex outcomes in speciation. It can inhibit speciation by homogenizing distinct populations through neutral diffusion [7], referred to as "breakdown" of reproductive isolation [8, 6]. When selection against hybrids limits neutral introgression between hybridizing populations, tension zones can form [9], with variable outcomes depending on the strength of selection. If selection is weak, tension zones can stabilize, thereby stalling speciation while maintaining genetic differentiation (e.g. [10]). If selection against hybrids is strong, reduced hybrid fitness can reinforce premating barriers (e.g. assortative mating), strengthening barriers to gene exchange and advancing speciation ("reinforcement" [11, 12]). Alternately, genetic differentiation can increase through adaptive introgression from one species to another [13], or unique admixture of new, hybrid species [14, 3, 15]. These phenomena are not mutually exclusive, and variation and combinations exist along the continuum. Ultimately, our interpretation of hybridization only captures "a single snapshot of a complex and continuously changing interaction" [2].

Homoploid hybrid speciation (hereafter referred to as hybrid speciation) and adaptive introgression are particularly important promoters of differentiation, as they can generate novel hybrid entities as well as the resources to fuel adaptive divergence of preexisting species [5, 2, 13]. Although both processes involve hybridization of distinct populations or species, hybrid speciation culminates with adaptive novel hybrid combinations forming a distinct and independent hybrid taxon [16, 3, 17]. In adaptive introgression, recombinant hybridization passes adaptive variation from one species to another, replacing less adaptive portions of the original genome, but maintaining the majority of that original genome [18, 19, 13]. Defining hybrid speciation thus necessitates identification of novel, hybrid traits or combinations that allow hybrids to be distinguished from parental taxa, regardless of the level of genetic admixture [2]. Identification of these traits can be difficult both theoretically and empirically, and is further complicated in systems exhibiting complex phylogenetic histories [20, 21], differential influences of parental taxa [15], and ancient or repeated hybridization [22]. However, complex systems that exhibit repeated hybridization at different temporal and spatial scales provide useful "natural laboratories" for understanding the ramifications of hybridization and downstream processes at an evolutionary scale [7, 23].

Swallowtail butterflies of the *Papilio machaon* species group (Lepidoptera: Papilionidae: the Old World swallowtails) provide a diverse model system in which to study hybridization. Species limits and systematic relationships have been notoriously difficult to resolve (e.g. [24, 25] in part due to a plethora of ecological races, color morphs, and incomplete reproductive barriers, as well as abundant natural hybridization [26, 27, 28]. Additionally, the genus as a whole has been influential in the development of many prominent theories in biology (speciation [29], the biological species concept [30], coevolution [31], mimicry [32], etc.), and has had a disproportionate influence on our understanding of the genetic and ecological dynamics of hybrid speciation [33, 34, 35]. In North America, six species are currently recognized within the *P. machaon* species complex: *Papilio machaon* (the only member to have a Holarctic distribution [26]), *P. polyxenes*, *P. zelicaon*, *P. brevicauda*, *P. joanae*, and *P. indra* [36]. *Papilio indra* is the only North American member of the species group to consistently have distinctive genitalia and adult wing pattern [28] and multiple genetic studies place it as the sister species to the rest of the clade [37, 38]. The five remaining North American species share many wing pattern characteristics, but can generally be separated into yellow- or black-morph species (Fig 1), although color polymorphism is widespread (e.g. [39, 40], also [26]). Larval hostplant use in the *P. machaon* group is confined to species of Asteraceae, Apiaceae (Umbelliferae), and Rutaceae, and while geographical specialization is the norm, uncommon species/hostplant pairings have been widely observed, suggesting that ecological differentiation is not strongly tied to larval hostplant constraints [26, 27].
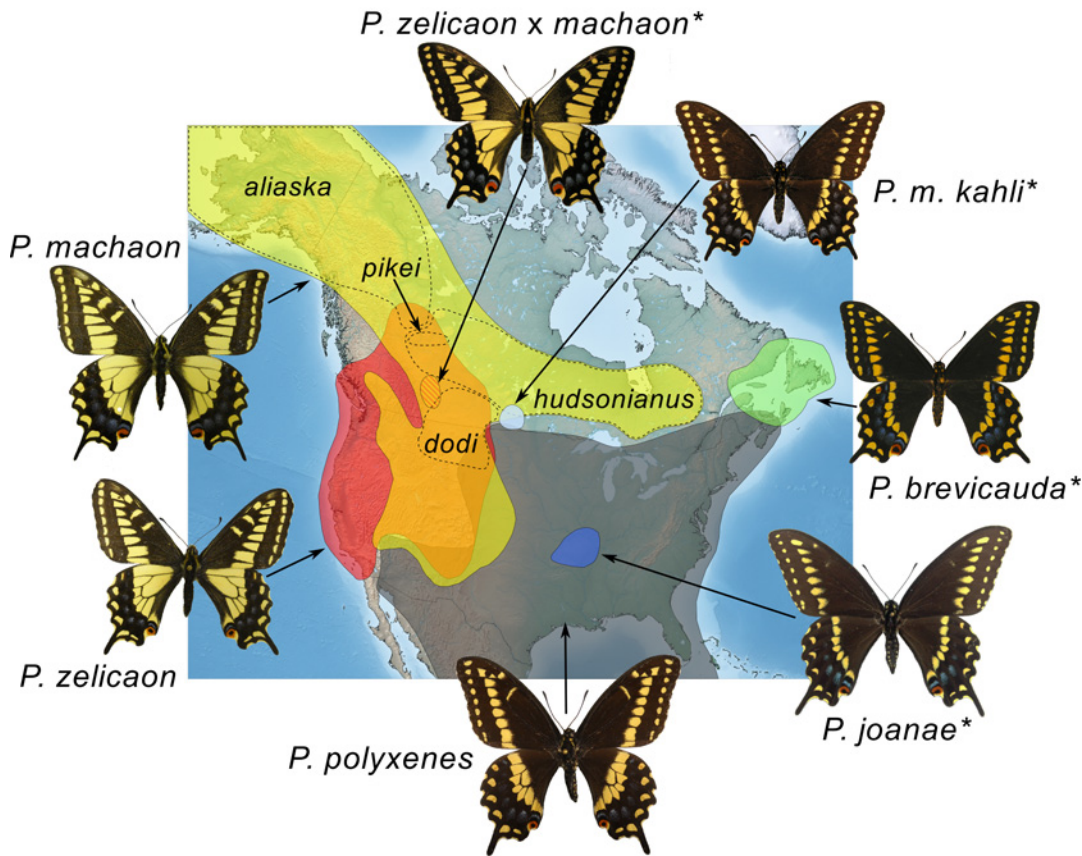
Fig 1. Generalized range map of current distributions of the *Papilio machaon* species complex in North America. Putative hybrid taxa are indicated with an asterisk. Dashed lines indicate approximate ranges of *P. machaon* subspecies pertinent to the current study. Map image: public domain from www.simplemappr.net, *Papilio joanae* holotype photograph: J. Tewell.

The impetus for the research reported here was the finding that *P. brevicauda* and *P. joanae*, although morphologically quite similar to *P. polyxenes* (Fig 1), share mitochondrial DNA (mtDNA) restriction-site patterns with *P. machaon* [28]. Both *P. brevicauda* and *P. joanae* are geographically separated from *P. machaon*. *Papilio brevicauda* is restricted to the Maritime Provinces of Canada, often frequenting bluffs near the sea [41, 42]. *Papilio joanae*, on the other hand, is an endemic species of closed woodland habitats of the Ozark plateau in Missouri [43, 44], and because of its restricted range has been classified as "vulnerable" by conservation associations [45]. Both of these taxa have variably been considered species in good standing (*P. brevicauda* [46], *P. joanae* [36]) or subspecies/races of *P. polyxenes* due to their morphological resemblance (*P. brevicauda* [47], *P. joanae* [43]). More recently, however, they have been described as close relatives of *P. machaon* based on the mtDNA findings of [28] [48, 42, 36]. Mitochondrial DNA haplotypes identical to those in *P. brevicauda* and *P. joanae* were also found by [28] in putative hybrids in southwestern Alberta (*P. machaon* x *P. zelicaon* [26]) and southern Manitoba (*P. machaon kahli* [39], or *P. machaon* x *P. polyxenes* after [26]). Both of these populations represent morphological intermediates between the putative parental taxa, and while morphology and allozymes have been studied in the SW Alberta populations [26, 49], those in southern Manitoba have received less attention. Following [26] and [36], respectively, we refer to the populations in southwestern Alberta as *P. zelicaon* x *machaon*, and those

in southern Manitoba as *P. machaon kahli*. "Parental taxa" refer to *P. machaon*, *P. polyxenes*, and *P. zelicaon*.

Despite the discordant mtDNA affinities of these taxa and their recognition as close relatives of *P. machaon*, no research has followed up on the hypothesis of ancient hybridization set out by [28]. Here we evaluate this putative hybridization across North America using mitochondrial and nuclear DNA sequences, microsatellite markers, molecular dating, morphological analysis of wing pattern, and assessment of ecological characteristics. Given the apparent evolutionary complexity of the group, we approach this evaluation using the preexisting taxonomic entities most prevalent in the literature (*sensu* [26, 36]) to guide our interpretation of these lineages, and ask to what extent each putative hybrid lineage (*P. brevicauda*, *P. joanae*, *P. m. kahli*, and *P. zelicaon* x *machaon*) exhibits characteristics of the putative parental species (*P. machaon*, *P. polyxenes*, and *P. zelicaon*). Using these data, we then evaluate whether significant support exists to classify each of these lineages as hybrid species, or if they appear to be the result of adaptive introgression. Following [2], we consider a hybrid species to be one that demonstrates novel characteristics or combinations that distinguish, and ideally, reproductively isolate the hybrid from *both* of its parents. Specifically, we use our genetic and morphological data to assess the genealogical discordance and distinctness of hybrid lineages, and focus on ecological separation to guide our inference of reproductive isolation (e.g. host plant, habitat, flight period, etc.). Genealogical and morphological distinctness, and reproductive isolation are considered to constitute significant evidence for the hypothesis of hybrid speciation.

Materials and Methods

Specimen selection and DNA extraction

Specimens were selected to represent the taxonomic breadth and diversity of the *P. machaon* species group in North America, as well as by the availability of DNA and wing vouchers. Species designations followed work on the same specimens or populations included in prior studies (e.g. [26, 28]). Morphology, geography, and in some cases larval host plant information were used to identify new specimens to species. We define species, in accordance with the genomic integrity species definition [50], as populations able to maintain their genomic integrity upon contact with each other, or allopatric populations with genetic divergence proportional to that found in closely related sister species. Slightly different subsets of specimens were used for each analysis due to DNA/wing quality limitations and voucher availability. Details are described below and species-specific sample sizes for each analysis are given in Table 1. All specimens used in this study, as well as associated locality and voucher deposition information, are given in Table A in S1 File (some geographic coordinates for private land and historic samples are not displayed). Collecting of recent material in provincial parks in Alberta and British Columbia was conducted under permit numbers 10–097 and 105180, respectively, and collecting on private land was done with the owner's permission. All freshly collected material was collected with an aerial net (for adults), or hand collected as larvae and reared to the adult stage, then killed and stored at -70°C.

Sequence data generated by [51, 52, 53, 54] was retrieved from GenBank for a number of specimens within the *P. machaon* species group (Table 1), as well as *P. xuthus*, which served as an outgroup in the phylogenetic analyses [53]. Data retrieved from these sources varied from the full COI/COII genes (including the intervening *tRNA-leucine* gene; 2288 bp) to partial COI (394 bp), and the full EF-1α gene (1010 bp). Additional mitochondrial sequence data was generated for this study from pre-existing DNA extractions representing the major mtDNA restriction-site haplotypes found in [28]. More recently collected butterflies were also sequenced for COI, and in these cases, adult butterflies (field collected or reared from field-

Table 1. Species and specimens examined in this study. Sample sizes for COI/COII sequence data (SEQ), microsatellites (MSAT), and morphological characters (MORPH). Number in parentheses in sequence data column corresponds to EF-1α.

Species	Subspecies	Seq	Msat	Morph
P. machaon	aliaska	13(1)	12	12
	bairdii	7(1)	4	7
	dodi	7	7	7
	gorganus	2(1)	2	2
	hippocrates	2(1)	2	3
	hudsonianus	7	6	6
	oregonius	1(1)	2	3
	pikei	8(1)	6	6
	kahli	3	4	3
P. polyxenes	americus	2(1)	3	4
	asterius	8(1)	11	9
	coloro	5(1)	6	6
	stabilis	1(1)	1	1
P. zelicaon		25(4)	24	19
P. zelicaon x machaon		23(1)	23	20
P. brevicauda	brevicauda	4	4	4
	gaspeensis	3(2)	2	3
P. joanae		6(2)	7	7
P. hospiton		1(1)	0	1
P. indra	indra	3(3)	3	3
	kaibabensis	1(1)	1	1
P. xuthus		1(1)	0	0
Total		133(25)	130	127

doi:10.1371/journal.pone.0141882.t001

collected larvae) were killed and stored at -70°C. Additional EF-1α data was also generated for a subset of specimens from each species, although due to the limited and discordant phylogenetic information content of this gene, sequencing was not pursued for the remainder of the specimens (see Results). Microsatellite analysis was attempted on all specimens for which sequence data was available, except for a handful of specimens for which no DNA remained (including the only available specimen of *P. hospiton*). All recent DNA extractions (for both sequence data and microsatellites) were carried out using Qiagen DNeasy® Blood & Tissue extraction kits (QIAGEN, Mississauga, Ontario, CAN) using leg or thoracic tissue.

Sequence data

Polymerase chain reactions (PCRs) were conducted in 50 µL reactions with a Biometra TGradient thermal cycler (Biometra, Goettingen, DE), including the following reagents: for COI/COII, 5 µL 10x PCR buffer (Promega, Madison, WI, USA), 3 µL of 25 mmoles/µL MgCl2 (Promega), 1 µL of 10 mmoles/µL dNTPs (Roche, Switzerland), 2 µL of each forward and reverse primer in 5 pmol/µL concentrations, 1 µL of 5 U/µL *Taq* polymerase (QIAGEN), 1 µL DNA, and 35.5 µL autoclaved Millipore water; for EF-1α, all reagent quantities were identical except for 2 µL of MgCl2 and 36.75 µL Millipore water. Reactions were conducted with a hot start (introducing *Taq* Polymerase after the initial 2 minute, 94°C denaturation period) followed by 35 cycles of 94°C for 30 seconds, 45°C for 30 seconds, and 72°C for 1 minute for COI/COII, and 35 cycles of 94°C for 30 seconds, 55°C for 1 minute, and 72°C for 1.5 minute for EF-1α. All reactions were finished with a 7-minute final extension at 72°C. Primers used in this study are

given in Table C in S1 File. PCR purification was conducted with either a Qiagen QIAquick® PCR purification kit or a Qiagen QIAEX II® agarose gel extraction kit (QIAGEN). Sequencing reactions were carried out in both directions using a DYEnamic™ ET terminator cycle sequencing kit (Amersham Pharmacia Botech, Cleveland, Ohio, USA), and either filtered through Sephadex-packed columns or ethanol precipitated before being dried, resuspended, and fractionated on either an ABI PRISM® 377 or 3730 automated DNA sequencer (Applied Biosystems, Foster City, California, USA). Sequences were aligned using Mesquite v2.75 [55] and ClustalW v2.0.12 [56] using default settings. Alignment quality was checked by eye, but major adjustments were not necessary due to the absence of indels and introns. For the EF-1α sequences, double peaks consistently observed in the electropherograms were assumed to be the result of heterozygotes, and were coded using IUPAC ambiguity codes. The ends of sequences were trimmed to facilitate collapsing strictly redundant haplotypes in MacClade v4.08a [57]. Although in some cases this removed variable characters from the matrix for COI/COII, overall topological patterns were not affected.

Phylogenetic analyses

Phylogenetic analyses were conducted using multiple optimality criteria to ensure that the choice of analytical method did not bias conclusions. Unweighted, unordered maximum parsimony (MP [58]) searches were conducted in PAUP* 4.0b10-x8 [59], with heuristic strategy of 1000 replicates of random sequence addition (holding 10 trees per replication), tree-bisection and reconnection branch swapping (TBR [60]) and no limit to the maximum number of trees retained per replication. Bootstrapping [61] with heuristic strategy of 100 replicates of random sequence addition (holding 100 trees per replication), TBR, and a maximum number of 100 trees (of minimum score 1) retained per replication, was conducted to test node support. To test for incongruence between data sets, Templeton tests [62] of data heterogeneity were conducted in PAUP* on a subset of taxa (specimens that had data for both COI/COII and EF1α) to identify if one data set could statistically reject the topology of the tree given by the other data set. In these tests, 50% majority rule consensus trees of the individual complete data sets (from MP searches) were constrained to the alternative data set and evaluated using the PSCORES command in PAUP*.

Maximum likelihood (ML) analyses were conducted in GARLI v0.951-GUI [63] applying models of evolution as predicted by the Akaike Information Criterion [64] in jModelTest 2.1.1 [65, 66]. The following models of evolution were used: extended COI/COII and complete EF-1α: Transition model (TIM) + I; complete COI/COII and extended EF-1α: TIM + Γ. All parameter values were specified in GARLI, and 100 bootstrap replicates were conducted.

Bayesian inference (BI) was conducted in MrBayes v3.2 [67]. Transition models of evolution are not applicable in MrBayes so were simplified, as in [68], as follows for these analyses: extended COI/COII: general time reversal (GTR [69]) + I + Γ; complete COI/COII and extended EF-1α: GTR + Γ; and complete EF-1α: Hasegawa, Kishino, and Yano model (HKY [70]) + I. Two million generations were run with trees being sampled from both runs every 100 generations, and default chain settings/temperatures. No priors were specified. The average standard deviation of split frequencies was observed during the run and the potential scale reduction factors were observed after the run to ensure that independent simulations were converging (values should approach zero and one, respectively). Burn-in trees were estimated visually for both runs in the log-likelihood overlay plot, and 25% of sampled trees were removed for burn-in. Posterior probability (clade credibility) values were calculated in MrBayes, and a 50% majority rule consensus tree was constructed in PAUP* after removing burn-in trees.

Microsatellite markers

Ten of 17 microsatellite loci developed by [71] were reliably amplified in all species of interest (Table B in S1 File). Reverse primers for six of these loci were "PIGtailed" to decrease non-template nucleotide addition that hinders genotyping [72]. Microsatellite amplification was conducted using universal fluorescently labeled M13 forward primers [73], and sequence-specific primers mixed in a ratio of 4:1 reverse primer: M13 tailed forward primer. PCR reactions were conducted in 15 µL volumes containing 1.5 µL 10x microsatellite PCR buffer, 1.5 µL of 25 mM $MgCl_2$, 0.3 µL dNTPs, 0.48 µL 4:1 sequence specific primer mix, 0.48 µL universal fluorescent-labeled M13 primer, 0.2 µL *Taq* DNA polymerase (Pickard Laboratory, University of Alberta), and 2.5 µL DNA, under the following cycling conditions: 10 min at 94°C, 38 cycles of [30 sec at 94°C, 45 sec at 56°C or 57°C (Table B in S1 File)], and 45 sec of 72°C, and followed by a final extension of 10 min at 72°C. Amplified fragments were run on an ABI Prism 3730 Analyzer (ABI), with a Genescan® LIZ-500 size standard, and genotyped using Genemapper® v4.0 (ABI). Descriptive statistics and measures of population differentiation (F_{ST}) were calculated using GENODIVE [74].

Individual-based clustering

Bayesian clustering of individual microsatellite data was conducted in STRUCTURE [75] using an admixture model and independent allele frequencies for all analyses. A burn-in period of 150,000 Markov chain Monte Carlo (MCMC) generations was followed by 500,000 generations for $k = 1$ through $k = 10$, with 10 iterations for each value of k. The most likely number of genetic clusters was calculated by evaluating the likelihood of the data (lnP(D|K) [75]) and Δk [76] with the program STRUCTURE HARVESTER [77]. CLUMPP v1.1.2 [78] was used to average replicate runs for each k value. STRUCTURE was also run using putative parental taxa as training sets, species determinations as "population" priors, and on a dataset including *P. indra*, the outgroup. Results using either training sets or population priors did not differ from analyses without these conditions, so the latter are presented here. Analyses including *P. indra* are provided in Fig. B in S1 File. Sub-structure was assessed in the overall analysis (by calculating ancestry for suboptimal Δk values), and by breaking the dataset up according to the overall $k = 2$ results where individuals with $\geq 70\%$ *machaon*-like ancestry were treated separately from remaining individuals. These methods produced similar assessments of substructure: The latter are focused on in the results section, as they provided clearer demarcation of substructure, and the former are provided in Fig. C in S1 File.

To investigate relatedness between clusters, we also conducted discriminant analysis of principal components (DAPC [79]), which submits genetic data to a principal component analysis (PCA) before conducting discriminant analysis (DA) on those principal components. This multivariate discriminant method does not attempt to minimize Hardy-Weinberg and gametic equilibrium (as does STRUCTURE [75]), and is therefore potentially more suited to this style of phylogenetically-oriented sampling. In maximizing between- and minimizing within-group variability [79], DAPC has also been shown to be more powerful and accurate with hierarchical relationships [80], which might be predicted in situations of hybridization and differential introgression. We implemented DAPC in R v3.0.1 [81] using *adegenet* v1.3.1 [82]. To provide comparison with STRUCTURE, the *find.clusters* function was used with default parameters, retaining all principal components (PCs), to find the ideal k value.

To visualize relationships between clusters using DAPC, the *optim.a.score* function was used to determine the optimal number of PCs to retain in the DA. In this function, 25 full simulations (parameter *smart* = FALSE) of a preliminary DAPC run (retaining the number of PCs corresponding to one-third the sample size of the run) are reiterated by the *optim.a.score*

function to determine the ideal number of PCs to retain. This optimal number of PCs was then used in the final DAPC. All discriminant functions were retained for all DAPCs. The *xvalDapc* function was also used as an alternative to *optim.a.score*, and presented consistently similar, though somewhat larger (approximately 10–15 additional PCs), values of the optimal number of PCs. Due to uneven sample sizes between clusters, and potential biases in the *xvalDapc* function in these cases, *optim.a.score*'s determination of the ideal number of PCs was used (though retaining *xvalDapc*'s ideal number of PCs did not change the overall clustering pattern). Additionally, *optim.a.score*'s smaller optimal number of PCs graphically clustered the groups with smaller samples sizes closer to the main groupings, allowing better determination of relationships. DAPC was also conducted on a dataset including *P. indra*, the outgroup, and those results are provided in Fig. D in S1 File.

Morphometrics

Six wing morphometric characters were used as in [26] (pg 208–209): A) extent of yellow scaling in cell Cu2, in anal margin of dorsal hindwing, B) shape of pupil in anal eyespot of dorsal hindwing, C) extent of black scales between blue and red portions of anal eyespot of dorsal hindwing, E) extent of yellow scales in basal half of disc of ventral forewing, F) extent of yellow scales of postmedian yellow band in apical cell of ventral forewing, and G) number of cells with orange patch in postmedian area of ventral hindwing, plus one. Right wings were used, unless characters were only visible on left wings. The quality and preserved tissues of 20–30 year old voucher specimens limited the amount of useable characters to those on the wing, and wing length characters were ignored due to large amounts of missing data associated with worn/tattered specimens; some specimens for which DNA-based data were available were too damaged to be scored (Table 1). Multiple correspondence analysis (MCA), an alternative to PCA for categorical variables, was conducted in R v3.0.1 [81] with the package FactoMineR [83].

Molecular dating

Molecular dating of the complete COI/COII data was implemented in BEAST v1.6.2 [84], using BEAUti to generate the associated *xml* input file. Due to the paucity of fossil papilionids (see [38]), secondary calibrations from previous molecular dating studies in the Papilionidae were used to calibrate the tree. Five calibration points (shown in Fig. A in S1 File), including the root calibration, were used from the soft-bound age estimates calculated by [38] (nodes 167, 168, 169, 171, and 172). These were applied using the *tmrca* prior with a uniform distribution. Additionally the *Site Model* prior was set to match that used for BI (GTR + Γ), and the *Clock Model* prior was estimated with a relaxed, uncorrelated lognormal clock [85]. Both the Yule [86] and Birth-Death [87] process speciation *Tree Priors* were used, but did not change the results appreciably; results from the Birth-Death process prior will be reported. Five independent runs of 100×10^6 generations, being sampled every 1×10^3 generations, were run and combined using LOGCOMBINER. To ensure proper parameter estimates, TRACER was used to check effective sample sizes. A burn-in of 25% of the trees from each run was removed before all combined trees were summarized in TREEANNOTATOR.

Results

DNA sequence properties

Sequence data was collected from 133 individuals representing 22 taxa (Table 1, GenBank accession numbers: Table A in S1 File). DNA alignments for the gene regions *cytochrome oxidase I/II* (COI/COII) and *elongation factor-1α* (EF-1α) were 2288 bp and 1010 bp respectively

Table 2. Summary results from maximum parsimony and maximum likelihood analyses. Extended and complete data sets are included.

Gene	Data set	Characters		MP				ML Score
		Inf.*	Total	# Trees	Score	CI	RI	
COI/COII	Extended	164	2288	15186400	463	0.793/0.668	0.895	-5369.1201
	Complete	24	306	28	71	0.817/0.675	0.882	-753.4862
EF-1α	Extended	16	1010	22976344	86	0.919/0.720	0.877	-1927.2094
	Complete	9	418	859	42	0.952/0.833	0.957	-831.0601

Abbreviations: CI: consistency index, RI: retention index, Inf. *: parsimony informative characters, MP: maximum parsimony, ML: maximum likelihood. For CI values, the first reported number includes uninformative characters, and the second excludes them.

doi:10.1371/journal.pone.0141882.t002

(Table 2), although most sequences were shorter than this total. To test whether missing data affected overall topology, all phylogenetic analyses were conducted with the "extended" dataset (including missing data) and with a "complete" dataset that included only shorter sequence regions present in all specimens. The shorter alignments consisted of 306 and 418 bp, for COI/COII and EF-1α respectively.

The COI/COII data included 133 specimens that were consolidated to 54 unique haplotypes in the complete dataset, and the EF-1α data included 27 specimens that contained 25 unique genotypes (Table 1, Table A in S1 File). Tests of topological incongruence concluded that the two data sets (COI/COII and EF-1α) were not homogenous. Constraining the abridged COI/COII data to the EF-1α topology supported incongruence between the datasets ($p = 0.0002$); reversing that constraint (i.e. constraining EF-1α data to COI/COII topology) did not ($p = 0.0588$), as could be expected since the EF-1α data contained relatively little variation, and hence phylogenetic information (Table 2), and had few resolved clades (see below).

Phylogenetic relationships

Maximum parsimony, ML, and BI analyses produced very similar, although not identical, topologies for each data set. Measures of branch support (MP and ML bootstrap and BI posterior probability) generally increased for phylogenies based on extended data for COI/COII, but decreased for EF-1α. Overall topology was not affected by the inclusion of regions of missing data. Bayesian 50% majority rule consensus trees are shown here for the extended datasets (Fig 2), and summary information for MP and ML are presented in Table 2.

COI/COII supported previously resolved relationships [26, 51, 53, 88, 38] between *P. machaon*, *P. hospiton*, *P. polyxenes*, and *P. zelicaon* (Fig 2B). The mitochondrial phylogeny failed to separate subspecies within *P. machaon* and *P. polyxenes* (Table A in S1 File) indicating high intraspecific variability or retained ancestral polymorphism within these species, which is discordant with patterns defined by morphology and geography. Complete fixation of one clade of *machaon*-like mtDNA was observed in *P. brevicauda*, *P. joanae*, and *P. m. kahli*, while *P. zelicaon* x *machaon* hybrid populations contained that clade as well as *machaon*- and *zelicaon*-like haplotypes more typical of the parental species. All putative hybrids except *P. m. kahli* showed within-group variability, and several haplotypes within the main hybrid clade were shared between species (e.g. hyb11 and hyb12). The main hybrid clade was monophyletic with regard to putative hybrid populations, except for the presence of two *P. machaon* haplotypes: hyb13 from a single *P. m. pikei* specimen collected in Alberta and hyb12 in several *P. m. hudsonianus* specimens collected in Manitoba and Québec (Fig 1). Interestingly, the sister lineage to the main hybrid clade is a highly supported, divergent (long branch lengths) clade
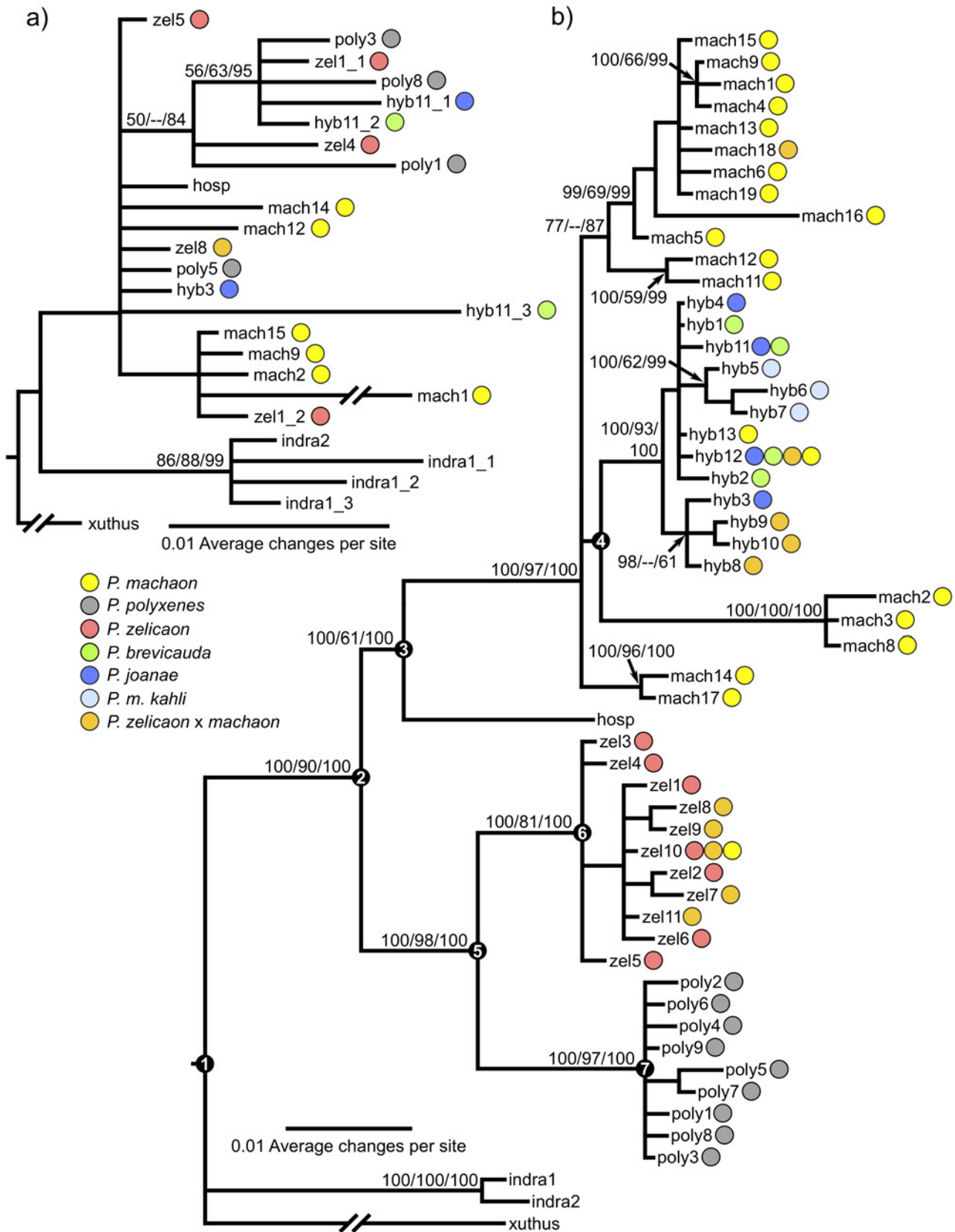
Fig 2. Fifty percent majority rule consensus trees constructed using Bayesian Inference. a) Extended EF-1α sequence data for 25 genotypes, and b) extended COI/COII sequence data for 54 haplotypes. Numbers above branches represent bootstrap support (if >50%) for maximum parsimony (first number), maximum likelihood (second number) and Bayesian posterior probability values (third number), for nodes that are consistently resolved among search criteria. Colors in circles correspond to the species possessing each haplotype. EF-1α genotype names reflect the corresponding COI/COII haplotype displayed by each specimen, with an added underscore and number indicating cases where specimens that shared COI/COII haplotypes had more than one different EF-1α genotype. Age estimates using COI/COII data for major nodes (numbered black circles) are provided in Table 4.

doi:10.1371/journal.pone.0141882.g002

belonging to two specimens of *P. m. aliaska* collected from northern British Columbia and Alaska (mach2 and mach3) and one specimen of *P. m. hudsonianus* collected from Manitoba (mach8). Additionally, two specimens identified as *P. m. aliaska* (based on morphology, flight period, and habitat) exhibited *zelicaon*-like mtDNA (zel10); upon further examination of these specimens, they showed several intermediate morphological characters between *P. m. aliaska* and *P. zelicaon*.

EF-1α sequences only supported the monophyly of *P. indra*, and gave no resolution for any other species within the group (Fig 2A). Due to the ambiguous phylogenetic information content of EF-1α and difficulty in consistently obtaining sequences from older DNA, comprehensive sequencing of all specimens was not pursued.

Microsatellite data

Ten microsatellite loci were genotyped for 130 specimens representing 20 taxa (Table 1), and had a total of 225 alleles with a range of 11–32 alleles per locus (Table B in S1 File). Observed heterozygosity ranged from 0.25 in *P. indra* to 0.67 in *P. zelicaon* x *machaon* and pairwise F_{ST} ranged from 0.010 between *P. zelicaon* and *P. zelicaon* x *machaon* to 0.600 between *P. m. kahli* and *P. indra* (Table 3). STRUCTURE predicted an ideal *k* value of *k* = 2, with clusters roughly corresponding to 1) *P. machaon* including *P. m. kahli*, and 2) *P. polyxenes*, *P. brevicauda*, *P. joanae* and *P. zelicaon* (Fig 3A); individuals of *P. zelicaon* x *machaon* were split between the two main clusters or, along with some individuals of *P. zelicaon*, were intermediate. Sub-structuring was present for both overall clusters. Within the *polyxenes*/*zelicaon*-like cluster, *k* = 2 separated *P. polyxenes* and *P. brevicauda* from *P. zelicaon* and *P. joanae* (Fig 3B), and *k* = 4 identified *P. brevicauda* as having a unique signature (Fig 3C). Within the *machaon*-like cluster, *k* = 5 was best supported, which clearly distinguished *P. m. kahli* and *P. m. pikei*, but showed variable ancestry for the other subspecies of *P. machaon* and individuals of *P. zelicaon* x *machaon* from SW Alberta (Fig 3D). Two individuals of *P. m. aliaska* were *zelicaon*-like or *zelicaon* x *machaon*-like throughout the STRUCTURE results (Fig 3); these were the same two individuals that exhibited *zelicaon*-like mtDNA (Fig 2B).

Table 3. F_{ST} and heterozygosity values based on microsatellite data. F_{ST} comparisons based on microsatellite data between species in bottom triangle, heterozygosity values within each species on diagonal (observed/expected).

|	polyx.	brevi.	joanae	zelicaon	zel x mac	kahli	mach.	indra
P. polyxenes	0.47/0.81	-	-	-	-	-	-	-
P. brevicauda	0.114	0.51/0.57	-	-	-	-	-	-
P. joanae	0.068	0.223	0.61/0.78	-	-	-	-	-
P. zelicaon	0.047	0.143	0.049	0.62/0.87	-	-	-	-
P. zelicaon x machaon	0.058	0.181	0.065	0.010	0.67/0.84	-	-	-
P. m. kahli	0.245	0.456	0.309	0.202	0.217	0.44/0.41	-	-
P. machaon	0.113	0.216	0.096	0.067	0.047	0.236	0.50/0.82	-
P. indra	0.281	0.481	0.375	0.239	0.291	0.600	0.288	0.25/0.43

doi:10.1371/journal.pone.0141882.t003
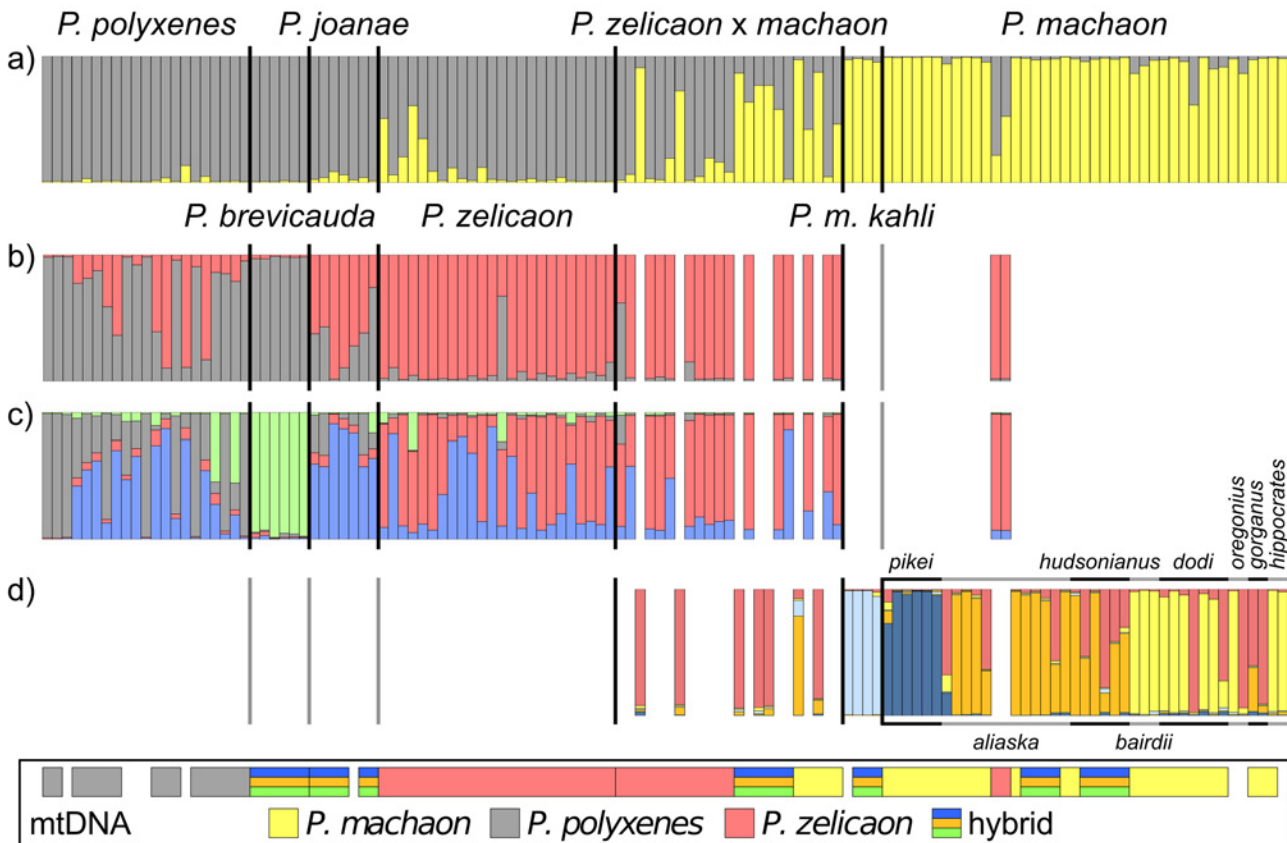
Fig 3. STRUCTURE results for microsatellite genotype data. a) *k* = 2 based on overall dataset; b) and c) *k* = 2 and *k* = 4, respectively, for the *polyxenes*/*zelicaon*-like cluster from the overall *k* = 2; and d) *k* = 5 for the *machaon*-like cluster from the overall *k* = 2 (individuals with ≥70% *P. machaon* ancestry). Inset indicates the major mtDNA clade (*P. machaon*, *P. polyxenes*, *P. zelicaon*, or the main hybrid clade within the *P. machaon* clade) for each individual (gaps indicate specimens genotyped for microsatellites that were not sequenced for COI/COII). Alternating black and grey bars above and below the *P. machaon* portion of d) indicate subspecies; from left to right: *P. m. pikei*, *aliaska*, *hudsonianus*, *bairdii*, *dodi*, *oregonius*, *gorganus*, and *hippocrates*.

doi:10.1371/journal.pone.0141882.g003

Discriminant analysis of principal components (DAPC) of microsatellite data optimally predicted *k* = 3 genetic clusters, corresponding to one group consisting of *P. machaon* and a few *P. zelicaon* x *machaon*, and two groups sharing all individuals of the other species and a small number of *P. machaon*. Graphically, there is no clear demarcation/separation between these groups; individuals instead fall along a gradient from *P. machaon* to *P. zelicaon* to *P. polyxenes* (Fig 4). *Papilio joanae* grouped between *P. polyxenes* and *P. zelicaon*, specimens of the *P. zelicaon* x *machaon* population were between their putative parental species, and *P. brevicauda* clustered closest to *P. polyxenes*, but with some separation from the main gradient. The contributions of individual discriminant functions are displayed as an inset in Fig 4, showing that the first function accounts for most of the variation. *Papilio machaon kahli* clusters among *P. machaon* on the first discriminant axis (the x-axis, discriminant function 1), but is separated from the main cluster along the second (the y-axis, discriminant function 2). This separation could be an artifact of small sample size capturing very little within-group variability, but generating additional "individuals" from random alleles observed in *P. m. kahli* maintains this overall pattern.

Considering the overall pattern of genetic clustering and observed intraspecific variation, STRUCTURE and DAPC produce similar results; namely, both analyses show optimal clustering of individuals that separates most *P. machaon* from *P. polyxenes*/*P. zelicaon*, although
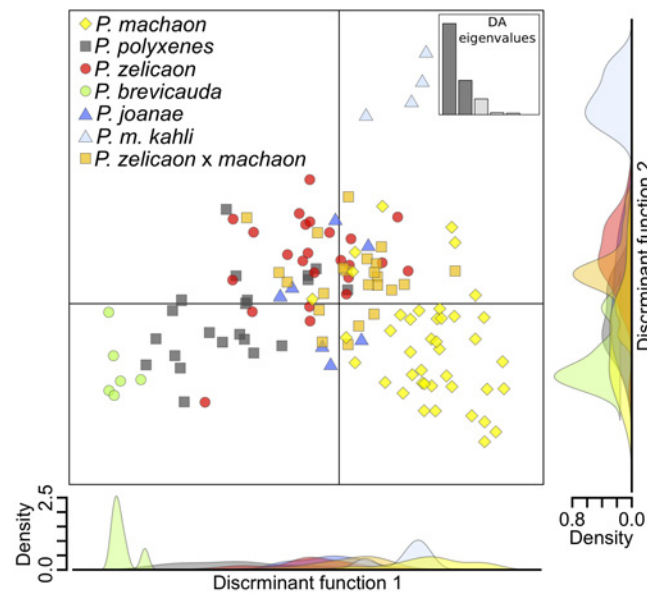
Fig 4. DAPC of putative hybrids and parental taxa generated from microsatellite data. Inset illustrates the relative contribution of individual discriminant functions (DFs) to overall variability, and the density plots of the two plotted DFs are shown outside of the main plot.

doi:10.1371/journal.pone.0141882.g004

STRUCTURE illustrates this pattern more explicitly. It is clear that despite their mitochondrial relationships, *P. brevicauda* and *P. joanae* share little nuclear resemblance with *P. machaon*. *Papilio machaon kahli* appears *machaon*-like at a broad level, but along with *P. brevicauda*, is genetically distinct at a finer scale. Finally, *P. zelicaon* x *machaon* shows signatures of high and variable admixture in both STRUCTURE and DAPC, which is exemplified by the occurrence of *machaon*-like, *zelicaon*-like, and hybrid-like individuals present within each of the main mtDNA clades ([Fig 3], inset).

Morphometrics

Six wing morphometric characters from [26] were scored for 127 specimens representing 21 taxa ([Table 1]), and analyzed with multiple correspondence analysis (MCA) ([Fig 5]). Dimension one of the MCA created a gradient between the parental species, stretching from "yellow morph" species (*P. machaon*) to "black morph" species (*P. polyxenes*); substantial variability was present, however, particularly in *P. machaon*. The putative hybrids showed less within-species variability when compared to the parental species, although smaller sample sizes likely contributed to this pattern. *Papilio brevicauda* and *P. joanae* grouped closely with *P. polyxenes*, which is unsurprising based on their overall appearance ([Fig 1]). *Papilio machaon kahli* was found on the periphery of the *P. polyxenes* cluster, extending towards *P. zelicaon* & *P. machaon*. Finally, *P. zelicaon* x *machaon* hybrids grouped closely within the *P. zelicaon*/*P. machaon* cluster, with one black morph ("nitra") individual extending towards the *P. polyxenes* region. Several individuals of *P. machaon* and *P. polyxenes* also occupy the intermediate space between the main groupings of their respective taxa (four *P. machaon* individuals to the left of x = -0.25, and three *P. polyxenes* individuals to the right of x = -0.4), and correspond to black morph *P. machaon bairdii* and yellow morph *P. polyxenes americus*.
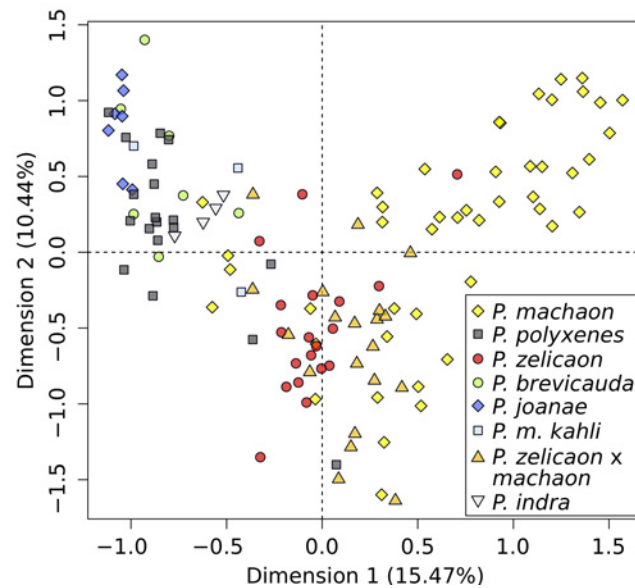
Fig 5. Multiple correspondence analysis of six morphological characters. Plotted using the *jitter* function to improve visualization of overlapping points (non-*jittered* results shown in Fig. E in S1 File).

Molecular dating

Molecular dating of the complete COI/COII data set resulted in a root age for the divergence of *P. xuthus* from the remaining species of 18.25 (15.29–23.38 95% confidence interval) million years (MY) before present (Fig 2, Table 4), which, as expected, is quite similar to the age of the respective node found by [38]. This leads to the age of divergence of the main hybrid haplotype clade of 1.60 (0.24–2.45) MY, or in the mid-Pleistocene. The overall topology of this analysis was consistent with that obtained from MrBayes, except the monophyly of the main hybrid group was further broken by the presence of two *P. machaon* haplotypes (Fig. A in S1 File). Age estimates for main nodes in Fig 2 are provided in Table 4, and the entire dated tree is shown in Fig. A in S1 File.

Discussion

Interspecific hybridization across the *P. machaon* species group in North America is supported by several mutually-reinforcing new lines of evidence. Using mtDNA sequence data, we identified a *machaon*-like lineage shared by four putative hybrid populations or species, confirming

Table 4. Age estimates and 95% confidence intervals of major nodes shown in Fig 2. All ages in millions of years.

Clade #	Age estimate	95% CI
1	18.25	15.29–23.38
2	7.07	5.02–9.01
3	5.12	3.77–7.01
4	1.60	0.24–2.45
5	3.46	0.51–6.71
6	1.51	0.08–3.73
7	1.46	0.07–3.67

earlier findings based on mtDNA restriction-site variation [28]. In contrast, nuclear markers and morphological characters exhibit variable admixture and intermediacy, ranging from signatures indistinguishable from *P. polyxenes* or *P. zelicaon* to those of a stable hybrid swarm. The one nuclear gene that was sequenced, EF-1α, showed only shared sequence variation among the major species in the group, although it confirmed *P. indra* as the closest outgroup. Hybridization among species would have been facilitated by the repeated glaciations of the Pleistocene, and illustrates the importance that hybridization can have in the evolutionary histories of entire species groups. We first discuss this phylogeographic hypothesis, and then consider the relative roles of hybrid speciation and adaptive introgression in the formation of each of these hybrid lineages.

Pleistocene origins and phylogeography of hybrid lineages

Molecular dating of COI/COII approximates the time of divergence of the main hybrid lineage as mid Pleistocene (Fig 2, Table 4). The use of a single marker and secondary calibrations make this a rough estimation (e.g. [89] but see [90] regarding the use of multiple secondary calibration points). Nonetheless, with confidence intervals of 1.1–2.8 MY, the initial hybridization events for all hybrid lineages can be confidently placed in the context of the repeated glaciations of the Pleistocene [91]. As [28] hypothesize, at these glacial maxima, the ranges of all three parental taxa would have been forced south, creating new regions of contact or sympatry between *P. machaon* and *P. polyxenes*/*P. zelicaon*, and facilitating hybridization (Fig 6). Subsequent glacial contraction likely separated sympatric populations, exposing hybrid remnants to differential introgression from their parental species.
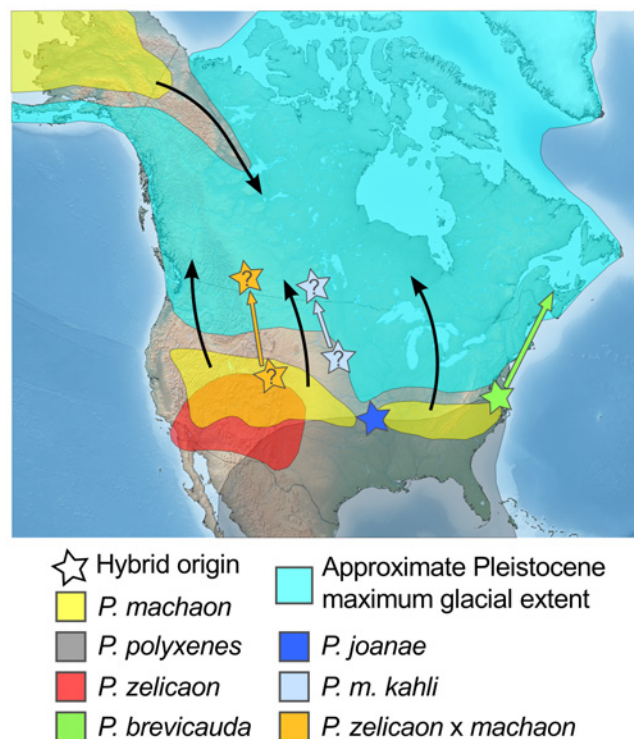

Fig 6. Approximate maximum glacial extent during the early- to mid-Pleistocene, hypothesized glacial refugia and hybrid origins. Arrows indicate hypothesized postglacial recolonization routes of hybrids (colored arrows), and parental species (black arrows). Hybrid origins with question marks indicate origins with less geographic certainty. Map image: public domain from www.simplemappr.net.

doi:10.1371/journal.pone.0141882.g006

Papilio machaon hudsonianus is the most easterly-distributed subspecies of *P. machaon* in North America [43] (Fig 1). The fact that *P. m. hudsonianus* mtDNA haplotypes (haplotype mach8) belong to the same mtDNA clade as the putative hybrid taxa supports a previously more widespread distribution for this mtDNA lineage. During or shortly after glacial maxima, *P. m. hudsonianus* from eastern North America would have been likely to come into contact with *P. polyxenes* in or near the current ranges of *P. joanae* and *P. brevicauda* (Fig 6). Interestingly, one specimen of *P. m. pikei* also displayed hybrid clade mtDNA (haplotype hyb13), and the sister clade to the main hybrid clade was found in individuals of both *P. m. hudsonianus* and *P. m. aliaska* (clade mach2/3/8). *Papilio machaon pikei* and *P. m. aliaska* have ranges parapatric or sympatric to that of *P. m. hudsonianus* in western Canada [43, 49] (Fig 1), and some gene flow between these three subspecies may be expected. However, both *P. m. aliaska* and *P. m. pikei* more commonly have typical *machaon*-like mtDNA haplotypes (e.g. haplotypes mach11, mach12, mach19) rather than those from the hybrid clade, suggesting that the presence of the hybrid clade haplotypes is due to relatively recent mitochondrial gene flow from *P. m. hudsonianus* to *P. m. aliaska* and *P. m. pikei*.

Two individuals that were identified as *P. m. aliaska*, based on general appearance, flight period, and habitat, exhibited *zelicaon*-like signatures for both mtDNA and microsatellites. Morphologically, these individuals clustered closer to *P. zelicaon* x *machaon* individuals, but exhibited a very *P. m. aliaska*-like overall appearance. They were collected alongside many typical *P. m. aliaska* as well as several individuals exhibiting hybrid-like mtDNA at a locality at the southern edge of the range of *P. m. aliaska*. These variable hybrid signatures, all observed on the same day at the same locality, illustrate the scale of evolutionary complexity that is observed in this group. Hybrid-like and more divergent *P. machaon* mtDNA signatures could be the remnants of an ancestral *P. machaon* population lineage once widespread in central and eastern Canada, which took part in hybridization at several locations. More geographically comprehensive sampling of *P. m. hudsonianus* and *P. m. aliaska* from the entirety of their ranges would shed further light on this hypothesis.

Alternative explanations for similar, and in some cases identical, *machaon*-like mtDNA haplotypes found in geographically disjunct putative hybrids are not supported by our data. The likelihood of this repeated geographic pattern arising by neutral chance (i.e. genetic drift [92]) is low, and although incomplete lineage sorting can generate phylogenetic patterns that mimic introgression, it would not be expected to leave any appreciable phylogeographic signal [93, 94]. *Wolbachia*, a bacterial endosymbiont that can cause cytonuclear discordance (e.g. [95]), has not been detected in the species group (Dupuis *personal observation*). Finally, if neutral processes such as drift or founders' effects were to fix introgressed haplotypes in putative hybrid populations, allelic (or haplotype) diversity would be expected to be low [96], which was not observed.

Hybrid speciation in the *P. machaon* group

Our data supports hybrid origins for four more-or-less distinct populations of the *P. machaon* species group. These lineages arose from similar situations, namely mating between *P. machaon* (likely *P. m. hudsonianus*) and *P. polyxenes*/*P. zelicaon*, but interestingly this has produced different outcomes in their current genetic composition and ecological associations. This repetition provides a unique framework to compare the outcomes of hybridization with regard to hybrid speciation and adaptive introgression. Both processes require hybridization between distinct biological entities, and generally result in some kind of cytonuclear discordance. While the distinction between hybrid speciation and adaptive introgression may seem arbitrary outside of speciation theory, biodiversity estimation and conservation generally rely

on species as a common unit of diversity. In this setting, it is important whether adaptive traits have introgressed from one species to another or two species have hybridized to create a third (e.g. [97, 98, 99]).

The key to distinguishing hybrid species from cases of adaptive introgression is to diagnose unique hybrid traits or combinations that facilitate reproductive isolation of hybrids from *both* parents [2]. Although genomic admixture and morphological intermediacy are conceptually intuitive indicators of potential hybridization, they can be difficult to tease apart from ancestral polymorphism and recent gene flow [13], and by themselves may not be enough to define a species as a homoploid hybrid [2]. This is especially true in cases of backcrossing to a single parent species [3], as would be expected with the *P. machaon* complex. Therefore, ecological evidence of reproductive isolation from both parents is often used in defining hybrid species (e.g. [100, 101]), although this isolation need not be absolute (e.g. [102, 34, 103]).

Here we focus on two criteria for defining a lineage as a hybrid species: 1) a hybrid species must have shared characteristics with both parents, but also some level of distinctiveness that facilitates diagnosis of the hybrids; and more importantly, 2) a hybrid species must have some novel characteristics that facilitate reproductive isolation from the parental species [2]. For the purposes of this study, meeting both of these criteria provides support for the lineage in question being a hybrid species; while meeting only one criteria (e.g. displaying genealogical discordance or introgression, but no characteristics of reproductive isolation) is insufficient evidence, and is more likely a scenario involving adaptive introgression in the history of the lineage. Although this framework greatly simplifies a conceptually complicated and difficult task [2], it is appropriate given the nature of this genetic data (non-genomic) and the paucity of ecological information for these butterflies, which are often difficult to find. To assist in this evaluation, we have compiled pertinent ecological traits known to facilitate reproductive isolation in Table D in S1 File, and have summarized this information, as well as our morphological and genetic conclusions, in Fig 7.

Papilio brevicauda. To lepidopterists familiar with swallowtails, *P. brevicauda* is a clearly diagnosable entity: its combination of *polyxenes*-like (black morph, orange undersides of hind wings) and *machaon*-like traits (short hindwing tails and shorter, rounded forewings) is distinctive and supports its widely-recognized species status. Only the *polyxenes*-like characters were included in our morphological analysis, due to the limited quality of voucher specimen wings and subsequent missing data, which can account for its indeterminate placement in the MCA. Genetically, *P. brevicauda* also displays characteristics of both *P. machaon* and *P. polyxenes*, as well as some degree of distinctiveness in nuclear markers at a fine-scale, supporting its consideration as a hybrid species. However, it appears to have no clear ecological separation from either parental species that would provide reproductive isolation from them (Fig 7) [43, 42, 104, 105], although one unknown in this ecological assessment is the eastern range limit of *P. m. hudsonianus*. Historic and potentially ongoing introgression of *P. m. hudsonianus* genes into a northern-adapted, coastal lineage of *P. polyxenes* could account for *P. brevicauda*'s morphological and genetic intermediacy, but it requires that that *machaon*-like mtDNA experienced a selective sweep through the population [106]. However, with the data at hand, the evidence for novelty of putative hybrid characteristics or reproductive isolation is not substantial, and so *P. brevicauda* does not fully qualify as a hybrid species. Comprehensive geographic sampling, particularly at the western and southern edges of the range of *P. brevicauda* (where it may be sympatric or parapatric with *P. m. hudsonianus* and *P. polyxenes*, respectively), should clarify its status.

Papilio joanae. Unlike *P. brevicauda*, *P. joanae* is often morphologically indistinguishable from *P. polyxenes asterius* [43, 45], and has only a handful of *machaon*-like traits. Except for mtDNA, its genetic characteristics are more *polyxenes*-like, although at fine scales it is
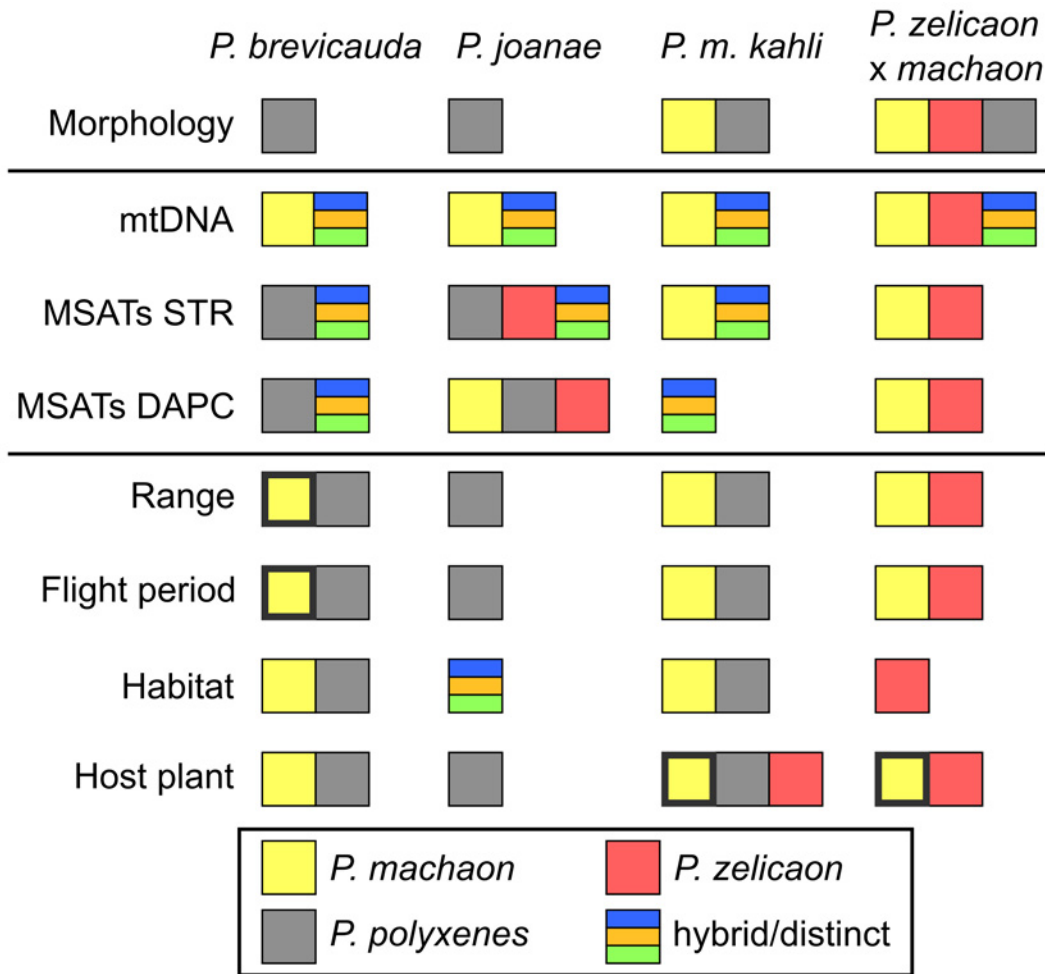
Fig 7. Summary of parental similarity and distinctiveness of four putative hybrid lineages with morphological, genetic, and ecological data.
Colored boxes denote similarity or clustering with parental species, or distinctness of the hybrid lineage. Morphology is based on clustering in morphological analysis (MCA). mtDNA is based on mitochondrial DNA clades (note that the unique hybrid clade is identified as both distinct and as *P. machaon*-like). MSATs STR is based on summarized ancestry at the finest subpopulation scale in STRUCTURE analyses, and MSATs DAPC on summarized clustering in DAPC analyses. Range shows sympatry or parapatry with parental species, flight period shows overlapping adult flight period, habitat is based on a general assessment of shared habitat type (forest openings, strict hilltopping behavior, under forest cover, etc.), and host plant shows shared larval host. Boxes with wide outlines indicate uncertainty in particular characteristics. Ecological information is presented in more detail in Table D in S1 File.

doi:10.1371/journal.pone.0141882.g007

somewhat *zelicaon*-like (Figs 3 and 4). Many of *P. joanae*'s ecological traits are shared with *P. p. asterius* [107, 43], except for its affinity for closed forests and cedar glades [107], rather than the open habitats (fields, exposed hilltops) that are used by the rest of the species group. This strict use of forest habitats nonetheless provides substantial separation of the two species, as *P. joanae* larvae are only found on hosts within forest habitats and *P. p. asterius* larvae are found only in open areas (Heitzman *personal communication*). Interestingly, *P. m. hudsonianus* is the only other North American member of the species group that frequents forest edges and shaded habitats, and even oviposits in shaded areas (Dupuis *personal observation*), although it uses different hosts than *P. joanae*.

The novelty of this *P. m. hudsonianus*-like ecological characteristic, which contributes to reproductive isolation of *P. joanae* from *P. polyxenes*, supports the hypothesis that *P. joanae* is indeed a hybrid species. This isolation is similar to ecological separation in several other North

American hybrid butterfly species [100, 34]. Adaptive introgression of *machaon*-like genes into *P. polyxenes* is also possible, but this alternative explanation is less likely. The nearest populations with similar mtDNA haplotypes are now over 1000 kilometers to the north. At finer scales, *P. joanae* displays *zelicaon*-like as well as *polyxenes*-like nuclear characteristics (Figs 3 and 4), suggesting older hybridization between *P. m. hudsonianus* and the common ancestor of *P. polyxenes* and *P. zelicaon*. For these reasons, we conclude that *P. joanae* is a homoploid hybrid species that is reproductively isolated from its parents via behavioral separation. This may be an important consideration for future conservation prioritization of *P. joanae* [45].

Papilio machaon kahli. *Papilio machaon kahli* is the most enigmatic of the putative hybrids considered here. Morphologically, it is intermediate between *P. machaon* and *P. polyxenes* (Fig 5), and lepidopterists have found it difficult to distinguish it from *P. polyxenes* based on overall appearance [108, 42]. Genetically, both mtDNA and nuclear DNA show ties to *P. machaon* (Figs 2 and 3A), but at a finer scale *P. m. kahli* is quite distinct from the rest of the species group (Figs 3D and 4). However, more samples are needed to elaborate this potential unique signature. We find little ecological support for any traits that would provide reproductive isolation from *P. machaon* or *P. polyxenes*; the use of its main host, *Zizia aptera*, may provide separation from *P. m. hudsonianus*, although the range of hosts used by the latter subspecies is unclear [26, 48]. Interestingly, in the past 100 years there may have been a decline in the presence of *kahli*-like individuals throughout the small range of this taxon, and an increase in the presence of *P. p. asterius* [26]. Unfortunately our study includes only one specimen collected more recently than 1990 (Table A in S1 File), so we cannot attest to the current status of that trend, but microsatellite data clustering shows little sign of *polyxenes*-like ancestry. Overall, we find no support for hybrid species status, and, based on the data at hand, we consider it most likely that *P. m. kahli* is a transitional population of *P. m. hudsonianus* experiencing adaptive introgression from *P. polyxenes*. The geographically limited range of this lineage was likely instrumental in its taxonomic recognition as a subspecies.

Papilio zelicaon x machaon. Populations of *P. zelicaon* x *machaon* in SW Alberta display a very different scenario compared to other potential hybrid taxa, in that hybrids are parapatric with both parental taxa, and all three mtDNA types are found in the same populations. Nuclear admixture is also variable among individuals, resembling early-generation hybrids (F1, F2, backcross) and both *P. machaon* and *P. zelicaon*. Morphological variation in these populations mirrors the nuclear admixture, although black ("nitra") morphs resembling *P. polyxenes* are also observed (Fig 7). Host plants and habitat preferences may provide hybrids with some ecological separation from *P. m. dodi*, which feeds strictly on *Artemesia dracunculus* in arid river valleys, but not from *P. zelicaon* [26]; adult hilltopping locations of *P. m. dodi* and *P. zelicaon* are often close enough to each other that both species can be found in the other's respective habitat. The southern range limit of *P. m. hudsonianus* is unclear, although *hudsonianus*-like specimens have been observed in southern Alberta in the vicinity of hybrid populations [49]. Whether these rare occurrences represent migrants from farther north or persisting populations is unknown. If they do represent the latter, then host choice may also foster ecological isolation between *P. m. hudsonianus* and *P. zelicaon* x *machaon* individuals.

The presence of mitochondrial, nuclear, and morphological intermediates at varying stages of evolutionary separation indicates that these populations represent a stable hybrid swarm [4, 109], and are far from being a distinct hybrid species. Predominant *zelicaon*-like ancestry is indicated with nuclear markers (Figs 3 and 4), but more comprehensive geographic sampling would clarify the situation. Interestingly, morphologically similar intermediates have been collected from the area since the early twentieth century before significant agricultural habitat changes [26], indicating that there was no anthropogenic influence on the initial formation of a hybrid population. Whether this situation will progress to hybrid speciation, as seen with *P.*

joanae, will depend on the development of mechanisms for ecological and reproductive isolation from both parental species.

Other lepidopteran hybrids

New World Lepidoptera have had disproportionate influence on the study of hybridization, due in part to their high diversity and general appeal, and these systems provide a rich foundation to compare to that of the *P. machaon* group. *Lycaeides* butterflies in western North America share many similarities with the *P. machaon* group, particularly in the multifarious nature of hybridization observed in the group [110]. Widespread historic admixture throughout *Lycaeides* is reminiscent of the *P. m. hudsonianus*-like signatures observed across North America in putative hybrid lineages. Additionally, ecological separation from parental taxa has allowed some hybrids to persist in novel habitats, akin to *P. joanae* [100]. *Heliconius* butterflies in Central and South America regularly hybridize (e.g. [111, 112, 113]), similarly to members of the *P. machaon* complex, and abundant ecological and genomic data has revealed promiscuous exchange of genes controlling protective color-pattern between hybridizing species [17]. However, contention regarding the hybrid origins of some *Heliconius* species [114] emphasizes the importance of addressing the fine differences between hybrid speciation and adaptive introgression. Within the Papilionidae, ecological and genomic data support the hybrid status of *Papilio appalachiensis* [33, 35, 115], where it is also clear that ecological separation has aided reproductive isolation from the parental species [34], again similarly to *P. joanae*. With growing insight into the potential for hybridization to encourage diversification and adaptation (e.g. [110, 17]), we are confident that continued work on the *P. machaon* group will facilitate this understanding, and support [46] assertion "that the Machaon-group provides some of the most suitable material ever investigated in animals for studying the process of speciation in detail."

Conclusions

Our results demonstrate a case of repeated reticulate evolution within a species complex of swallowtail butterflies in North America. We have documented three geographically separated cases of cytonuclear discordance where mtDNA is completely fixed in hybrid lineages (*P. brevicauda*, *P. joanae*, and *P. m. kahli*), and one case where both hybrid and parental mitochondrial haplotypes occur (*P. zelicaon* x *machaon*). Excluding the parental-like mtDNA haplotypes found in *P. zelicaon* x *machaon* populations, all of the mtDNA of hybrids is likely derived from a single *machaon*-like lineage resembling *P. m. hudsonianus*, despite geographic separation of these populations from *P. machaon* and from each other. Nuclear markers show variable signatures ranging from almost completely paternal characteristics, to high levels of admixture and potentially unique hybrid signatures; morphological characters also show variable levels of intermediacy. The divergences of these hybrid lineages were dated to the mid-Pleistocene, indicating an important role for glacial refugia in their formation. Despite similar hybrid origins (*P. machaon* hybridizing with *P. polyxenes*/*P. zelicaon*), these lineages have followed distinct evolutionary trajectories leading to diverse outcomes, from hybrid speciation (*P. joanae*) and potential adaptive introgression (*P. brevicauda* and *P. m. kahli*) to stable hybrid swarms (*P. zelicaon* x *machaon*). These results add to a growing recognition of the evolutionary importance and complexity of hybridization in generating biodiversity.

Supporting Information

S1 File. Supplementary file. Contains: Fig. A in S1 File. Molecular dated tree based on COI/COII data secondarily calibrated with four nodes; Fig. B in S1 File. STRUCTURE results for all

microsatellite data, including *P. indra*; Fig. C in S1 File. STRUCTURE results for microsatellite genotype data for the overall dataset, including substructure; Fig. D in S1 File. DAPC for all microsatellite data, including *P.indra*; Fig. E in S1 File. Morphometric MCA without using the *jitter* function to separate overlapping points; Table A in S1 File. Specimen information; Table B in S1 File. Microsatellite loci used in this study; Table C in S1 File. Primers for new sequences used in this study; and, Table D in S1 File. Summary of morphological and ecological information pertinent to the species included in this study.
(PDF)

S2 File. Microsatellite data.
(STR)

S3 File. Morphological data.
(TXT)

Acknowledgments

We thank B. Brunet, C. Davis, M. Dear, L. Lumley, B. Mori, C. Nargang, J. Janes, J. Pinzon, and K. Silva-Brandão for technical advice and assistance, G. Anweiler, B. Brunet, G. Hilchie, B. Mori, and C. Whitehouse for help in specimen collection, and two anonymous reviewers and the editor, F. Cotterill, for help in improving this manuscript.

Author Contributions

Conceived and designed the experiments: JRD FAHS. Performed the experiments: JRD. Analyzed the data: JRD. Contributed reagents/materials/analysis tools: JRD FAHS. Wrote the paper: JRD FAHS.

References

1. Wagner WH Jr. Biosystematics and evolutionary noise. Taxon. 1970; 19: 146–151.
2. Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, et al. Hybridization and speciation. J Evol Biol. 2013; 26: 229–246. doi: 10.1111/j.1420-9101.2012.02599.x PMID: 23323997
3. Mallet J. Hybrid speciation. Nature. 2007; 446: 279–283. PMID: 17361174
4. Nolte AW, Tautz D. Understanding the onset of hybrid speciation. Trends in Genetics. 2010; 26: 54–58. doi: 10.1016/j.tig.2009.12.001 PMID: 20044166
5. Butlin R, Debelle A, Claudius K, Snook RR, Beukeboom LW, Castillo Cajas RF, et al. What do we need to know about speciation. Trends Ecol Evol. 2012; 27: 27–39. doi: 10.1016/j.tree.2011.09.002 PMID: 21978464
6. Harrison RG. The language of speciation. Evolution. 2012; 66: 3643–3657. doi: 10.1111/j.1558-5646.2012.01785.x PMID: 23206125
7. Hewitt GM. Hybrid zones—natural laboratories for evolutionary studies. Trends Ecol Evol. 1988; 3: 158–167. doi: 10.1016/0169-5347(88)90033-X PMID: 21227192
8. Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Mol Ecol. 2006; 15: 343–355. PMID: 16448405
9. Barton NH, Hewitt GM. Analysis of hybrid zones. Ann Rev Ecol Syst. 1985; 16: 113–148.
10. Ruegg K. Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. Evolution. 2008; 62: 452–466. PMID: 18039327
11. Servedio MR, Noor MAF. The role of reinforcement in speciation: theory and data. Ann Rev Ecol Evol Syst. 2003; 34: 339–364.
12. Via S. Natural selection in action during speciation. Proc Natl Acad Sci USA. 2009; 106: 9939–9946. doi: 10.1073/pnas.0901397106 PMID: 19528641

13. Hedrick PW. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. Mol Ecol. 2013; 22: 4606–4618. doi: 10.1111/mec. 12415 PMID: 23906376

14. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, et al. Major ecological transitions in wild sunflowers facilitated by hybridization. Science. 2003; 301: 1211–1216. PMID: 12907807

15. Jiggins CD, Salazar C, Linares M, Mavarez J. Hybrid trait speciation and *Heliconius* butterflies. Philos Trans R Soc Lond B Biol Sci. 2008; 363: 3047–3054. doi: 10.1098/rstb.2008.0065 PMID: 18579480

16. Gross BL, Rieseberg LH. The ecological genetics of homoploid hybrid speciation J Heredity. 2005; 96: 241–252.

17. Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, et al. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature. 2012; 487: 94–98. doi: 10.1038/nature11041 PMID: 22722851

18. Whitney KD, Randell RA, Rieseberg LH. Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. New Phytol. 2010; 187: 230–239. doi: 10.1111/j.1469-8137.2010.03234.x PMID: 20345635

19. Pardo-Diaz C, Salazar C, Baxter SW, Merot C, Figueiredo-Ready W, Joron M, et al. Adaptive introgression across species boundaries in *Heliconius* butterflies. PLoS Genet. 2012; 8: e1002752. doi: 10.1371/journal.pgen.1002752 PMID: 22737081

20. McGuire JA, Linkem CW, Koo MS, Hutchison DW, Lappin AK, Orange DI, et al. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution. 2007; 61: 2879–2897. PMID: 17941840

21. Melo-Ferreira J, Boursot P, Carneiro M, Esteves PJ, Farelo L, Alves PC. Recurrent introgression of mitochondrial DNA among hares (*Lepus* spp.) revealed by species-tree inference and coalescent simulations. Syst Biol. 2012; 61: 367–381. doi: 10.1093/sysbio/syr114 PMID: 22201159

22. Genner MJ, Turner GF. Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. Mol Biol Evol. 2012; 29: 195–206. doi: 10.1093/molbev/msr183 PMID: 22114359

23. Buggs RJA. Empirical study of hybrid zone movement. Heredity. 2007; 99: 301–312. PMID: 17611495

24. Ae SA. The phylogeny of some *Papilio* species based on interspecific hybridization data. Syst Entomol. 1979; 4: 1–16.

25. Clarke CA, Larsen TB. Speciation problems in the *Papilio machaon* group of butterflies (Lepidoptera: Papilionidae). Syst Entomol. 1986; 11: 175–181.

26. Sperling FAH. Evolution of the *Papilio machaon* species group in western Canada (Lepidoptera: Papilionidae). Quaestiones Entomologica. 1987; 23: 198–315.

27. Sperling FAH. Natural hybrids of *Papilio* (Insecta: Lepidoptera): poor taxonomy or interesting evolutionary problem? Can J Zool. 1990; 68: 1790–1799.

28. Sperling FAH, Harrison RG. Mitochondrial DNA variation within and between species of the *Papilio machaon* group of swallowtail butterflies. Evolution. 1994; 48: 408–422.

29. Mallet J. Poulton, Wallace and Jordan: how discoveries in *Papilio* butterflies led to a new species concept 100 years ago. System Biodivers. 2004; 1: 441–452.

30. Mayr E. Karl Jordan's contribution to current concepts in systematic and evolution. Transactions of the Royal Entomological Society of London. 1955; 107: 45–66.

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

32. Brower JVZ. Experimental studies of mimicry in some North American butterflies: II. *Battus philenor* and *Papilio troilus*, *P. polyxenes* and *P. glaucus*. Evolution. 1958; 12: 123–136.

33. Scriber JM, Ording GJ. Ecological speciation without host plant specialization; possible origins of a recently described cryptic *Papilio* species. Entomol Exp Appl. 2005; 115: 247–263.

34. Kunte K, Shea C, Aardema ML, Scriber JM, Juenger TE, Gilbert LE, et al. Sex chromosome mosaicism and hybrid speciation among tiger swallowtail butterflies. PLoS Genet. 2011; 7: e1002274. doi: 10.1371/journal.pgen.1002274 PMID: 21931567

35. Zhang W, Kunte K, Kronforst MR. Genome-wide characterization of adaptation and speciation in tiger swallowtail butterflies using *de novo* transcriptome assemblies. Genome Biol Evol. 2013; 5: 1233–1245. doi: 10.1093/gbe/evt090 PMID: 23737327

36. Pelham J. Catalogue of the butterflies of the United States and Canada. J Res Lepid. 2008; 40: xiv–658.

37. Emmel JF, Emmel TC. Genetic relationships of *Papilio indra* and *Papilio polyxenes*. J Res Lepid. 1964; 3: 157–158.

38. Condamine FL, Sperling FAH, Wahlberg N, Rasplus J-Y, Kergoat GJ. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. Ecol Lett. 2012; 15: 267–277. doi: 10.1111/j.1461-0248.2011.01737.x PMID: 22251895

39. Chermock FH, Chermock RL. Two new races of Papilios from Manitoba. Bulletin of the Southern California Academy of Sciences. 1937; 36: 11–13.

40. Remington CL. A new sibling *Papilio* from the Rocky Mountains, with genetic and biological notes (Insecta, Lepidoptera). Postilla. 1968; 119: 1–40.

41. Morris RF. Butterflies and moths of Newfoundland and Labrador: the macrolepidoptera. Quebec: Canadian Government Publishing Centre; 1980.

42. Layberry RA, Hall PW, Lafontaine JD. The butterflies of Canada. Toronto: University of Toronto Press; 1998.

43. Scott JA. The butterflies of North America: a natural history and field guide. Standford: Stanford University Press; 1986.

44. Allen RT. Insect endemism in the interior highlands of North America. Fla Entomol. 1990; 73: 539–569.

45. Schweitzer DF, Minno MC, Wagner DL. Rare, declining, and poorly known butterflies and moths (Lepidoptera) of forests and woodlands in the eastern United States. Washington DC: U.S. Forest Service, Forest Health Technology Enterprise Team; 2011.

46. Clarke CA, Sheppard PM. A preliminary report on the genetics of the *machaon* group of swallowtail butterflies. Evolution. 1955; 9: 182–201.

47. Rothschild W, Jordan K. A revision of the American Papilios. Novitates Zoologicae. 1906; 13: 411–752.

48. Tyler HA, Brown KS, Wilson KH. Swallowtail butterflies of the Americas: a study in biological dynamics, ecological diversity, biosystematics, and conservation. Gainsville, FL: Scientific Publishers Inc.; 1994.

49. Bird CD, Hilchie GJ, Kondla NG, Pike EM, Sperling FAH. Alberta butterflies. Edmonton: Provincial Museum of Alberta; 1995.

50. Sperling FAH. Butterfly molecular systematics: from species definitions to higher-level phylogenies. In Boggs CL, Watt WB, Ehrlich PR, editors. Butteflies: Ecology and Evolution Taking Flight. Chicago: University of Chicago Press; 2003. pp. 431–458.

51. Caterino MS, Sperling FAH. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. Mol Phylogenent Evol. 1999; 11: 122–137.

52. Reed RD, Sperling FAH. Interaction of process partitions in phylogenetic analysis: an example from the swallowtail butterfly genus *Papilio*. Mol Biol Evol. 1999; 16: 286–297. PMID: 10028294

53. Zakharov EV, Caterino MS, Sperling FAH. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). Syst Biol. 2004; 53: 193–215. PMID: 15205049

54. Bromilow SM, Sperling FAH. Phylogeographic signal variation in mitochondrial DNA among geographically isolated grassland butterflies. J Biogeogr. 2011; 38: 299–310.

55. Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis. Version 2.75. 2011. Available: http://mesquiteproject.org.

56. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23: 2947–2948. PMID: 17846036

57. Maddison DR, Maddison WP. MacClade 4: analysis of phylogeny and character evolution. Version 4.08a. 2005. Available: http://macclade.org.

58. Fitch WM. Toward defining the course of evolution: minimum change for a specific tree topology. Syst Zool. 1971; 20: 406–416.

59. Swofford DL. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sunderland MA: Sinauer Associates; 2002.

60. Swofford DL, Olsen GJ. Phylogenetic reconstruction. In Hillis DM, Moritz C, editors. Molecular Systematics. Sunderland MA: Sinauer Associates; 1990. pp. 411–501.

61. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985; 39: 783–791.

62. Templeton AR. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evolution. 1983; 37: 221–244.

63. Zwickl DJ. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Thesis, The University of Texas at Austin. 2006. Available: http://repositories.lib.utexas.edu/handle/2152/2666.

64. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974; 19: 716–723.

65. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012; 9: 722.

66. Guindon S, Gascuel O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst Biol. 2003; 52: 696–704. PMID: 14530136

67. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2011; 61: 539–542.

68. Zakharov EV, Lobo NF, Hellmann JJ. Introgression as a likely cause of mtDNA paraphyly in two allopatric skippers (Lepidoptera: Hesperiidae). Heredity. 2009; 102: 590–599. doi: 10.1038/hdy.2009.26 PMID: 19293835

69. Tavaré S. Some probabilistic and statistical problems on the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences. 1986; 17: 57–86.

70. Hasegawa M, Kishino H, Yano T. Dating the human-ape split by a molecular clock of mitochondrial DNA. J Mol Evol. 1985; 22: 160–174. PMID: 3934395

71. Zakharov EV, Hellman JJ. Characterization of 17 polymorphic microsatellite loci in the Anise swallowtail, *Papilio zelicaon* (Lepidoptera: Papilionidae), and their amplification in related species. Mol Ecol Notes. 2007; 7: 144–146.

72. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by taq DNA polymerase: primer modifications that facilitate genotyping. BioTechniques. 1996; 20: 1004–1010. PMID: 8780871

73. Schuelke M. An economic method for the fluorescent labeling of PCR fragments: a poor man's approach to genotyping for research and high-throughput diagnostics. Nat Biotechnol. 2000; 18: 233–234. PMID: 10657137

74. Meirmans PG, Van Tienderen PH. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes. 2004; 4: 792–794.

75. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155: 945–959. PMID: 10835412

76. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005; 14: 2611–2620. PMID: 15969739

77. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012; 4: 359–361.

78. Jakobson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007; 23: 1801–1806. PMID: 17485429

79. Jombart T, Devillard S, Balloux F. Discriminate analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010; 11: 94. doi: 10.1186/1471-2156-11-94 PMID: 20950446

80. Kanno Y, Vokoun JC, Letcher BH. Fine-scale population structure and riverscape genetics of brook trout (*Salvelinus fontinalis*) distributed continuously along headwater channel networks. Mol Ecol. 2011; 20: 3711–3729. doi: 10.1111/j.1365-294X.2011.05210.x PMID: 21819470

81. R Development Core Team: *R version 3.0.1*. Vienna: R Foundation for Statistical Computing; 2013. Available: http://www.R-project.org/.

82. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24: 1403–1405. doi: 10.1093/bioinformatics/btn129 PMID: 18397895

83. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008; 25: 1–18.

84. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7: 214. PMID: 17996036

85. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006; 4: e88. PMID: 16683862

86. Yule GU. A mathematical theory of evolution: based on the conclusions of Dr. J.C. Willis. Philos Trans R Soc Lond B Biol Sci. 1924; 213: 21–87.

87. Gernhard T. The conditioned reconstruction process. J Theor Biol. 2008; 253: 769–778. doi: 10.1016/j.jtbi.2008.04.005 PMID: 18538793

88. Simonsen TJ, Zakharov EV, Djernaes M, Cotton AM, Vane-Wright RI, Sperling FAH. Phylogenetics and divergence times of Papilionidae (Lepidoptera) with special reference to the enigmatic genera *Teinopalpus* and *Meandrusa*. Cladistics. 2011; 27: 113–137.

89. Graur D, Martin W. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. TRENDS Genet. 2004; 20: 80–86. PMID: 14746989

90. Forest F. Calibrating the tree of life: fossils, molecules and evolutionary timescales. Ann Bot. 2009; 104: 789–794. doi: 10.1093/aob/mcp192 PMID: 19666901

91. Mickelson DM, Colgan PM. The southern Laurentide ice sheet. In Gillespie AR, Porter SC, Atware BF, editors. Developments in Quaternary Science volume 1: The Quaternary Period in the United States. Amsterdam Elsevier; 2004: pp. 1–16.

92. Ballard JWO, Kreitman M. Is mitochondrial DNA a strictly neutral marker? Trends Ecol Evol. 1995; 10: 485–488. PMID: 21237113

93. Funk DJ, Omland KE. Species-level paraphyly and polyphyly: frequency, causes and consequences, with insights from animal mitochondrial DNA. Annu Rev Ecol Evol Syst. 2003; 34: 397–423.

94. McKay BD, Zink RM. The causes of mitochondrial DNA gene tree paraphyly in birds. Mol Phylogenet Evol. 2009; 54: 647–650. doi: 10.1016/j.ympev.2009.08.024 PMID: 19716428

95. Jiggins FM. Male-killing Wolbachia and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. Genetics. 2003; 164: 5–12. PMID: 12750316

96. Wilson CC, Bernatchez L. The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). Mol Ecol. 1998; 7: 127–132.

97. DeMarias BD, Dowling TE, Douglas ME, Minckley WL, Marsh PC. Origin of *Gila seminude* (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. Proc Natl Acad Sci USA. 1992; 89: 2747–2751. PMID: 1557380

98. Haig SM, Mullins TD, Forsman ED, Trail PW, Wennerberg L. Genetic identification of spotted owls, barred owls, and their hybrids: legal implications of hybrid identity. Conserv Biol. 2004; 18: 1347–1357.

99. Monzón J, Kays R, Dykhuizen DE. Assessment of coyote-wolf-dog admixture using ancestry-informative diagnostic SNPs. Mol Ecol. 2014; 23: 182–197. doi: 10.1111/mec.12570 PMID: 24148003

100. Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC. Homoploid hybrid speciation in an extreme habitat. Science. 2006; 314: 1923–1925. PMID: 17138866

101. Schwarz D, Shoemaker KD, Botteri NL, McPheron BA. A novel preference for an invasive plant as a mechanism of animal hybrid speciation. Evolution. 2007; 61: 245–256 PMID: 17348936

102. Hermansen JS, Sæther SA, Elgvin TO, Borge T, Hjelle E, Sætre G-P. Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. Mol Ecol. 2011; 20: 3812–3822. doi: 10.1111/j.1365-294X.2011.05183.x PMID: 21771138

103. Stemshorn KC, Reed FA, Nolte AW, Tautz D. Rapid formation of distinct hybrid lineages after secondary contact of two fish species (*Cottus* sp.). Mol Ecol. 2011; 20: 1475–1491. doi: 10.1111/j.1365-294X.2010.04997.x PMID: 21251111

104. Ferguson, DC. The Lepidoptera of Nova Scotia. Part I: Macrolepidoptera. Nova Scotia: Nova Scotia Museum of Science Bulletin no. I; 1955.

105. Scriber JM. Segregation of F-2 interspecific hybrid growth performance and wing color patterns relative to parental species in the *Papilio machaon* species group (Lepidoptera: Papilionidae). Insect Sci. 2007; 14: 389–400.

106. Bazin E, Glémin S, Galtier N. Population size does not influence mitochondrial genetic diversity in animals. Science. 2006; 312: 570–572. PMID: 16645093

107. Heitzman JR. A new species of *Papilio* from the eastern United States (Papilionidae). J Res Lepid. 1973; 12: 1–10.

108. Klassen P, Westwood AR, Preston WB, McKillpp WB. The butterflies of Manitoba: Manitoba Museum of Man and Nature; 1989.

109. Latch EK, Kierepka EM, Heffelfinger JR, Rhodes OE Jr. Hybrid swarm between divergent lineages of mule deer (*Odocoileus hemionus*). Mol Ecol. 2011; 20: 5265–5279. doi: 10.1111/j.1365-294X.2011.05349.x PMID: 22066874

110. Gompert Z, Lucas LK, Buerkle CA, Forister ML, Fordyce JA, Nice CC. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. Mol Ecol. 2014; 23: 4555–4573.

111. Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, Lianres M. Speciation by hybridization in *Heliconius* butterflies. Nature. 2006; 441: 868–871. PMID: 16778888

112. Mallet J, Beltrán M, Neukirchen W, Linares M. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. BMC Evol Biol. 2007; 7: 28. PMID: 17319954

113. Nadeau NJ, Martin SH, Kozak KM, Salazar C, Dasmahapatra KK, Davey JW, et al. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Mol Ecol. 2012; 22: 814–826. doi: 10.1111/j.1365-294X.2012.05730.x PMID: 22924870

114. Brower AVZ. Hybrid speciation in *Heliconius* butterflies? A review and critique of the evidence. Genetica. 2011; 139: 589–609. doi: 10.1007/s10709-010-9530-4 PMID: 21113790

115. Cong Q, Borek D, Otwinowski Z, Grishin NV. Tiger swallowtail genome reveals mechanisms for speciation and caterpillar chemical defense. Cell Reports. 2015; 10: 910–919.