Received: 18 February 2023 | Accepted: 20 April 2023

DOI: 10.1111/syen.12595

ORIGINAL ARTICLE

Systematic Entomology

Royal Entomological Society

Phylogenomics resolves major relationships of *Catocala* underwing moths

Nicholas T. Homziak[1,2] | Caroline G. Storer[1] | Lawrence F. Gall[3] | Robert J. Borth[4] | Akito Y. Kawahara[1]

[1]McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, Florida, USA

[2]Entomology and Nematology Department, University of Florida, Gainesville, Florida, USA

[3]Entomology Division, Yale Peabody Museum, New Haven, Connecticut, USA

[4]Lepidoptera Biodiversity, LLC, Venice, Florida, USA

Correspondence
Akito Y. Kawahara, McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, FL, USA.
Email: kawahara@flmnh.ufl.edu

Abstract

Underwing moths in the genus *Catocala* Schrank are among the most charismatic of Lepidoptera. *Catocala* is also one of the most diverse genera worldwide in the speciose family Erebidae, but a phylogenetic framework for the genus is lacking. Here we reconstruct the first comprehensive molecular phylogeny for the genus based on 685 anchored hybrid enrichment loci sampled from 161 *Catocala* species (99 Nearctic, 62 Palearctic), four species of *Ulotrichopus* Wallengren and 33 outgroups. Phylogenetic analysis unambiguously recovers *Catocala* and *Catocala* + *Ulotrichopus* as monophyletic with strong support and resolves many backbone relationships within *Catocala*. Our results confirm the classification of previously proposed taxonomic subgroups of *Catocala*, including seven based on recent molecular/morphological evidence, and ten based on early twentieth-century morphological research. Mapping of larval host plant use onto the tree shows Fabaceae to be the likely ancestral host plant family for *Catocala* and *Catocala* + *Ulotrichopus*. There appear to have been at least 18 independent larval host plant shifts to nine plant families, the most common shift being from Fabaceae to Fagaceae. Larval host plant use has likely played an important role in the evolutionary history of *Catocala*, with several rapid diversification events propelled by shifts to novel larval host plants, particularly in the North American *Catocala* fauna.

KEYWORDS
Erebidae, host plant evolution, systematics

INTRODUCTION

Underwing moths in the genus *Catocala* Schrank are among the most recognizable and taxonomically diverse moths in the family Erebidae, the latter of which includes approximately 25,000 described species worldwide (van Nieukerken et al., 2011). This iconic genus has attracted keen interest from lepidopterists for centuries (e.g., Holland, 1903; Lees & Zilli, 2019; Sargent, 1976) and has been the focus of substantial systematic research (see e.g., Barnes & McDunnough, 1918; Gall & Hawks, 2010; Ishizuka, 2011 and references therein). Adult *Catocala* moths are noted for their diverse morphological patterns, especially the dorsal hindwings, which have large and boldly coloured bands that vary from yellow, orange, red and pink to white or blue (Ishizuka, 2011; Sargent, 1976). *Catocala* currently includes 268 species that are distributed across the globe (103 Nearctic, 165 Palearctic; there are no Holarctic species). In the Nearctic, *Catocala* is the fourth most speciose genus among the 747 genera in Noctuoidea, and the most species rich among the 268 genera of Erebidae (see Zahiri et al., 2017). Comparable Palearctic treatments likewise support the species-rich nature of *Catocala* in that region (e.g., Fibiger & Hacker, 2005; Goater et al., 2003; Kononenko, 2010).

As one of the most diverse genera of Noctuoidea, it is of particular interest to understand what factors played a role in the diversification and speciation of *Catocala*. For example, larvae of *Catocala* are known to feed on just ten families of broad-leaved trees and shrubs and are monophagous to only modestly oligophagous (Gall, 1991a; Ishizuka, 2011; Sargent, 1978). Accordingly, larval host plant preference has been thought to be phylogenetically conserved (see e.g., Gall, 1991a; Ishizuka et al., 2011, 2015), but the lack of a comprehensive phylogeny for the genus has prevented detailed analysis.

In his treatise on the world noctuid fauna, Hampson (1913: 3–9) used adult morphological characters to attempt to classify the 109 genera in the subfamily Catocalinae. This classification relied heavily on the presence and absence of tibial spines, and Hampson (1913: vii-ix) postulated that the Afrotropical/Oriental genera *Audea* Walker and *Ulotrichopus* Wallengren were closely related to *Catocala* sensu lato. Hampson subdivided *Catocala* into four genera: *Catabapta* Hulst, *Catocala* Schrank, *Ephesia* Hübner and *Mormonia* Hübner. Many Palearctic workers adopted this arrangement, but Barnes and McDunnough (1918) outlined the shortcomings of the Hampsonian genera and maintained the single genus *Catocala*, which all Nearctic workers followed. Barnes & McDunnough also showed that characters from eggs, larvae, genitalia and adults could be used for species identification and classification, and erected 20 infrageneric "Groups" that comprised presumptively monophyletic sets of species.

Mitter and Silverfine (1988) conducted a comprehensive cladistic analysis of *Catocala* and allied genera, based on 55 morphological characters. Their study hypothesized that *Catocala* may be paraphyletic with respect to *Ulotrichopus*. A subsequent cladistic analysis of catocalines by Kühne (2005) proposed six autapomorphies for the Catocalini and established the current concept for the tribe. Under Kühne's classification, Catocalini consists of *Archaeopilocornus* Kühne, *Audea* Walker, *Catocala*, *Crypsotidia* Rothschild, *Hypotacha* Hampson, *Tachosa* Walker and *Ulotrichopus*. Kühne (2005) was unable to resolve the relationship between *Catocala* and *Ulotrichopus* but also suggested that the former was likely paraphyletic with respect to the latter.

In the past decade, nine *Catocala* species groups have been proposed as monophyletic based on phylogenetic analyses that combined adult morphological and molecular (Cytochrome c Oxidase subunit I [COI]) data. Five of these species groups comprise only Nearctic taxa and corroborate and/or refine the assessments of Barnes and McDunnough (1918), abbreviated hereon as B&McD. These five are the Rosaceae-feeding *grynea* species group (=B&McD Group XVII; Kons & Borth, 2015a), the Fabaceae-feeding *nuptialis* species group (=B&McD Group XIII; Kons & Borth, 2016), the Fagaceae-feeding *andromache* and *delilah* species groups (=B&McD Group XII; Borth & Kons, 2016; Hawks, 2010), and the Ericaceae-feeding *andromedae* species group (=B&McD Group XV; Kons & Borth, 2017). Three other species groups comprise only Palearctic taxa feeding (or presumed to feed) on Fagaceae. These include the *intacta* species group (Borth, Kons, & Saldaitis, 2017), the *naganoi* species group (Kons et al., 2017) and the *dissimilis* species group (Kons et al., 2016). The ninth species group includes both Nearctic and Palearctic taxa in the Salicaceae-feeding *nupta* species group (=B&McD Group X; Borth, Kons, Saldaitis, & Gall, 2017).

Two recent studies by Ishizuka et al. (2011, 2015) applied molecular phylogenetic approaches to examine relationships of *Catocala* species. Ishizuka et al. (2011) focused mainly on the Japanese and mainland Eurasian faunas, sampling 33 species and a single mitochondrial gene, ND5. Ishizuka et al. (2015) subsequently sampled the ITS2 and 28S genes across 144 species of *Catocala* from Africa, Asia, Europe and North America and mapped hindwing colour (binned as black, orange or red) onto their resulting tree. Despite the reasonably broad taxon sampling, these two molecular studies were limited to three or fewer loci, and results yielded low branch support at many nodes, especially at deeper parts of the tree. However, Ishizuka et al. (2015, p. 159) identified two groups with strong support among the *Catocala* they sampled and made two proposals about larval host plant associations. First, a set of 12 Nearctic species with red hindwings that feed on Salicaceae share a common ancestor with the Eurasian Salicaceae-feeding *Catocala nupta* (Linnaeus). Second, a set of 19 Nearctic species that feed on Juglandaceae share a common ancestor (with closest relatives perhaps being the tropical Southeast Asian *Ulotrichopus macula* (Hampson) and/or *Ulotrichopus sumatrensis* Prout).

The present study investigates *Catocala* relationships based on a taxon set of 198 species (161 *Catocala*, four *Ulotrichopus* and 33 outgroups) and 685 loci. This data set includes over half of the described world *Catocala* species and has more than 300 times the gene sampling of any previous phylogenetic study of the genus. We test the monophyly of *Catocala* and its previously suggested species groups, examine relationships between *Catocala* and *Ulotrichopus*, and conduct an ancestral state reconstruction analysis of larval host preference to commence an investigation of host plant evolution in *Catocala*.

MATERIALS AND METHODS

Taxon sampling

We initially sampled 210 taxa to build a phylogeny, representing 170 *Catocala* species (including three variants), four *Ulotrichopus* and 33 primarily erebid outgroup species. DNA sequences of outgroup taxa were obtained from Homziak et al. (2019) and Kawahara et al. (2019). Tissues for DNA isolation (Table S1) were sampled from specimens collected directly into 95% ethanol or from dried museum specimens. Most alcohol-preserved samples were obtained from moths collected at night using UV light or bait and stored at −80°C in the McGuire Center for Lepidoptera and Biodiversity (MGCL), Florida Museum of Natural History, University of Florida, Gainesville, FL, USA. The right forewing and hindwing from these specimens were removed and vouchered, following the methods of Cho et al. (2016) to facilitate visual identification. Dried specimens were obtained from papered material held in the MGCL collection, or from pinned specimens at MGCL and the Yale Peabody Museum.

DNA extraction and sequence capture

DNA was extracted from all specimens using the OmniPrep Genomic DNA Extraction Kit (G-Biosciences, St. Louis, MO, USA). For alcohol-

preserved specimens, the extraction protocol was modified following Breinholt et al. (2018). DNA extraction procedures for legs of dried specimens followed St Laurent et al. (2018). Extractions from abdomens utilized the semi-nondestructive approach of Hamilton et al. (2019). Abdomens >20 years old followed additional protocols adapted from Hundsdoerfer and Kitching (2010), which precede a standard genitalia dissection. Abdomens were removed from dried specimens with sterile forceps and placed with the anterior end up in a microcentrifuge tube, and 500 μL of a 1:50 ratio of proteinase K to lysis buffer solution was subsequently added to each (Hamilton et al., 2019). Abdomens were lysed at 56°C for 2–4 hours (Hamilton et al., 2019). Additional steps were adapted from Hundsdoerfer and Kitching (2010) to ensure more thorough digestion of abdominal tissues. Fine Iris surgical scissors were used to cut along the left pleural membrane as far as the seventh abdominal segment following standard Lepidoptera dissection protocol (Hardwick, 1950). Afterwards, 50 μL of proteinase K was injected using a micropipette to fill the lateral incision and cover the internal tissues at the top of the abdomen, then left to incubate overnight for 16–18 hours (Hundsdoerfer & Kitching, 2010). DNA concentration was measured using the Qubit dsDNA Broad Range Assay Kit in a Qubit 2.0 fluorometer (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA). Extract quality (e.g., DNA fragmentation) was evaluated visually using agarose gel electrophoresis. Extracts with a minimum concentration of 10 μg/mL were sent to RAPiD Genomics (Gainesville, FL, USA) for library preparation and sequencing.

Anchored hybrid enrichment (Lemmon et al., 2012) was used to obtain DNA sequence data for phylogenetic analysis. We applied the Lep1 Agilent Custom SureSelect probe kit (Breinholt et al., 2018) which targets 855 loci that have been demonstrated to work well to resolve relationships of Lepidoptera at multiple taxonomic levels (e.g., Breinholt et al., 2018; Johns et al., 2018; St Laurent et al., 2018; Yang et al., 2021). Sequence cleaning and phylogenetic analyses were conducted on the University of Florida HiPerGator High Performance Computing Cluster (http://www.hpc.ufl.edu/). Raw reads for each taxon were filtered for quality using TRIMGALORE! v.0.4.0 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Loci for each sample were assembled from the filtered reads using iterative baited assembly (Breinholt et al., 2018) which employs USEARCH (Edgar, 2010) and BRIDGER (Chang et al., 2015). Sequences were aligned using MAFFT v.7.245 (Katoh & Standley, 2013) prior to using NCBI BLASTN (Camacho et al., 2009) to check orthology of probe regions by comparison to the reference *Bombyx mori* (L.) genome (Xia et al., 2004). Consensus sequences from isoforms were generated using FASconCAT-G (Kück & Longo, 2014). Contamination and duplicates were removed using USEARCH and custom python scripts, and sequences were aligned in MAFFT. Lastly, probe regions were manually checked and fit to the correct reading frame, removing one or two nucleotide positions at the beginning and end of each sequence when necessary.

Phylogenetic analyses

Loci with sequence data for less than 75% of the species in the dataset were removed. Similarly, 9 taxa with less than 200 recovered loci were excluded from the dataset to minimize interference from missing data (Breinholt et al., 2018). DNA sequences were analysed using both concatenation and coalescent-based phylogenetic methods to account for potential differences in gene histories. The concatenated dataset was partitioned by codon position using PartitionFinder 2 (Lanfear et al., 2017), with the rcluster algorithm (Lanfear et al., 2014) set to search the top 1000 subsets with the command –rcluster-max 1000 and –rcluster-percent 10. Tree reconstruction from the concatenated nucleotide and amino acid datasets used IQ-TREE 2.0 (Minh et al., 2020), with branch support assessed using both the Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) (Guindon et al., 2010), and ultrafast bootstrap approximation (UFBoot) (Hoang et al., 2018; Minh et al., 2013). Coalescent-based gene-tree reconstruction was conducted in ASTRAL-III (version 5.7.5) with default parameters (Mirarab et al., 2014; Zhang et al., 2018), and branch support assessed using local posterior probabilities (LPP). We did not conduct a dating analysis for this study because fossils do not exist for *Catocala* and its close relatives (Sohn et al., 2015), and it was felt that a future study with greater outgroup sampling is needed to properly date the tree.

Topological hypothesis testing

A four-cluster likelihood mapping (FcLM) analysis (Strimmer & von Haeseler, 1997) was used to assess the phylogenetic placement of certain taxa in our tree, as an alternative to traditional branch support metrics. For the FcLM analyses, we tested whether: (a) *Ulotrichopus* is monophyletic; or whether (b) *U. macula* is more closely related to *Catocala* than other *Ulotrichopus* we examined. The four clusters chosen were: (1) *U. macula*; (2) *Catocala* sensu stricto; (3) African *Ulotrichopus* and (4) outgroups. The FcLM analysis was implemented in IQ-TREE using the same partitions and substitution models as the maximum likelihood (ML) analysis of DNA data. We specified 2000 random quartets to be drawn using the -lmap 2000 command.

Evolution of larval host plant associations

Larval host data were gathered from authoritative recent literature (Borth & Kons, 2016; Dubatolov & Kosterin, 2000; Gall, 1991a, 1991b, 1991c; Hawks, 2010; Hsu et al., 2021; Kons et al., 2017; Kons & Borth, 2015a; Kons & Borth, 2015b; Kons & Borth, 2016; Kroon, 1999; Sargent, 1976; Staude et al., 2016, 2020; Wagner et al., 2011) and from ongoing long-term research by LFG. Known host plant families and dominant host plant genera were recorded for all sampled species (see Table S2). We used SIMMAP in the R package Phytools (Revell, 2012) for ancestral state reconstruction. Larval host plant data were coded into a matrix of prior probabilities for each tip at the host plant family level (with two refinements at the genus level in Fagaceae and Juglandaceae) as input for analysis. Tips with missing host data were assigned equal prior probabilities for each host plant family. We specified symmetrical rates for forward and reverse character state transitions, and 1000 rounds of stochastic character

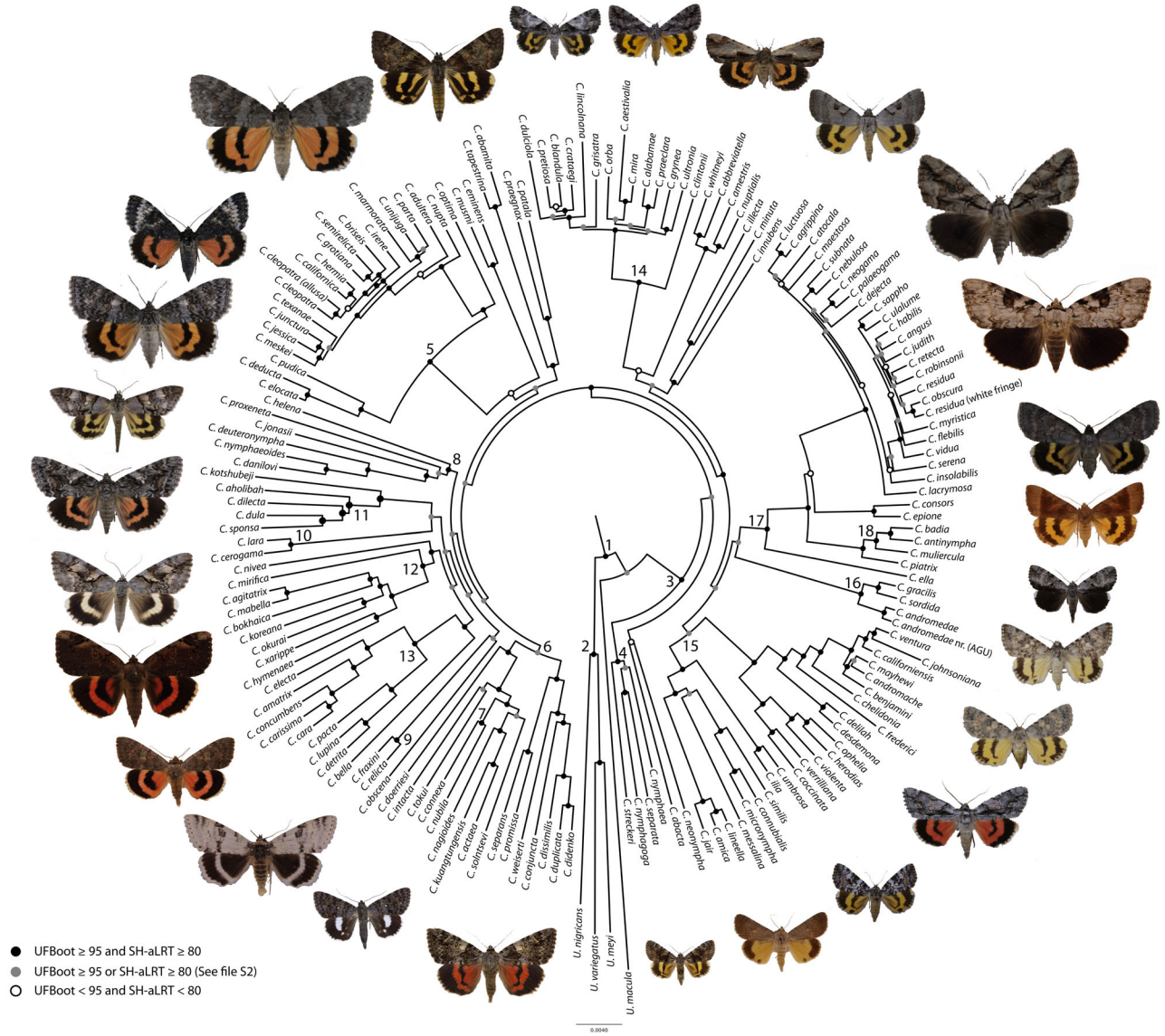

FIGURE 1 Maximum likelihood tree of 161 *Catocala* and 4 *Ulotrichopus* species based on 685 loci. Node support values were calculated using SH-like approximate likelihood ratio test (SH-aLRT), and ultrafast bootstrap approximation (UFBoot). Black dots represent nodes where UFBoot ≥95 and SH-aLRT ≥80 (threshold values for strong support from each metric). Grey dots represent nodes where either UFBoot ≥95 UFBoot or SH-aLRT ≥80. Empty dots indicate UFBoot <95 and SH-aLRT <80. Printed scores at each node and outgroup relationships are shown in file S2.

mapping to estimate posterior probabilities of character states at ancestral nodes and tips with missing data.

RESULTS

Taxon sampling and sequence capture

The final data matrix included 198 species (161 *Catocala* [including three variants], four *Ulotrichopus* and 33 outgroups) after exclusion of samples that were deemed unusable for phylogenetic analysis. Of the 168 ingroup samples, 88 were extracted from alcohol-preserved tissues and 80 were extracted from the abdomens or legs of pinned museum specimens. Ages of pinned museum specimens included in the final data matrix ranged from 3 months to 87 years, with a median age of 13 years at the time of DNA extraction. Detailed information for specimens extracted can be found in Table S1.

Anchored hybrid enrichment probes captured 685 loci with at least 75% taxon coverage across the 855 loci targeted by the Lep1 probe set. The full length of the concatenated probe-region sequences was 159,984 nucleotide sites with an average locus length of 233 nucleotide base pairs (standard deviation = 122.5). Raw sequence reads, individual FASTA alignments for probe regions of loci used in analyses, and the concatenated FASTA supermatrix are available at: 10.5061/dryad.n2z34tn1v.

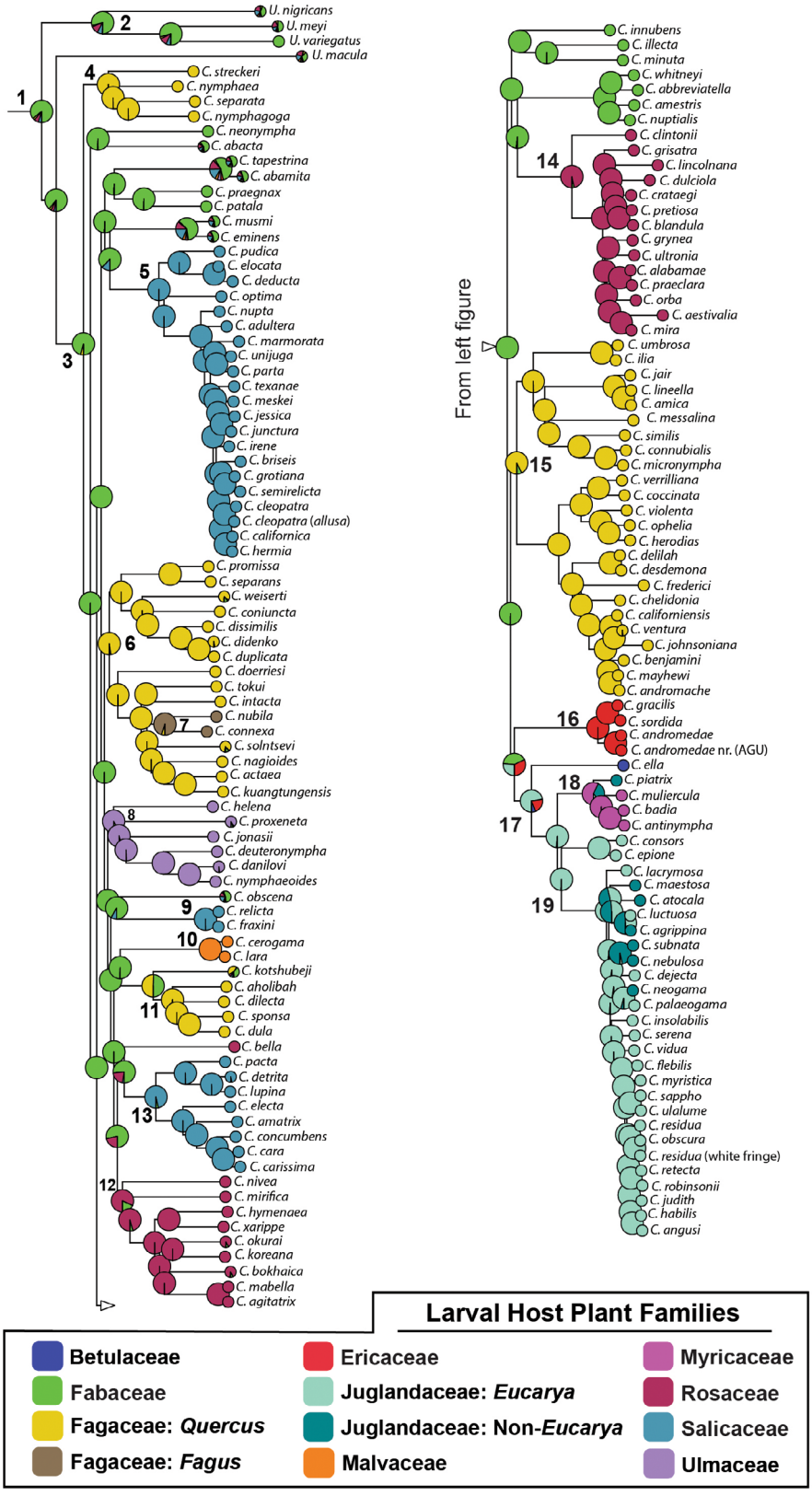

FIGURE 2 Ancestral state reconstruction of larval host plant family mapped to the maximum likelihood nucleotide tree. Green = Fabaceae, Brown = Fagaceae: *Fagus*, Yellow = Fagaceae: *Quercus*, Medium Blue = Salicaceae, Dark Blue = Betulaceae, Purple = Ulmaceae, Orange = Malvaceae, Maroon = Rosaceae, Red = Ericaceae, Teal = Juglandaceae: Non-*Eucarya*, Aquamarine = Juglandaceae: *Eucarya*, Magenta = Myricaceae. Node probabilities are shown in Table S3.

Phylogenetic analyses

Maximum likelihood trees of the concatenated nucleotide dataset (Figures 1, S1, S2) and amino acid dataset (Figures S3, S4) recovered strong support for the monophyly of *Catocala* + *Ulotrichopus* (UFBoot/SH-aLRT = 100/100; node 1), African *Ulotrichopus* (node 2) and *Catocala* (node 3; Figures 1, 2). Similarly, ASTRAL analyses recovered strong support for the monophyly of *Catocala* + *Ulotrichopus* (LPP = 1), African *Ulotrichopus* (LPP = 1) and *Catocala* (LPP = 1; Figures S5, S6). The nucleotide analysis placed the Asian *U. macula* as sister to *Catocala* with moderate support (UFBoot/SH-aLRT = 88/91.4). This relationship was also recovered in the ASTRAL analysis but with low support (LPP = 0.73), while the amino acid analysis placed *U. macula* as sister to the African *Ulotrichopus* clade with moderate support (UFBoot/SH-aLRT = 89/94.6). We did not find support for a paraphyletic relationship between *Catocala* and *Ulotrichopus* as indicated by Mitter and Silverfine (1988) and Kühne (2005). Our results suggest that *Catocala* and African *Ulotrichopus* may be monophyletic sister taxa.

Both ML and ASTRAL analyses recovered strong support for many clades within *Catocala*. Topological conflict between trees appeared largely limited to shallow nodes with short internode lengths within well-supported clades or deeper nodes with low branch support. Our results independently recovered seven of the nine recently proposed *Catocala* species groups as monophyletic, including the Fagaceae-feeding *andromache* (Borth & Kons, 2016), Ericaceae-feeding *andromedae* (Kons & Borth, 2017), Fagaceae-feeding *delilah* (Borth & Kons, 2016), Fagaceae-feeding *dissimilis* (Kons & Borth, 2016), Rosaceae-feeding *grynea* (Kons & Borth, 2015a), Salicaceae-feeding *nupta* (Borth, Kons, Saldaitis, & Gall, 2017) and Fabaceae-feeding *nuptialis* species groups (Kons & Borth, 2016). We were unable to test the monophyly of the Fagaceae-feeding *intacta* (Borth, Kons, Saldaitis, 2017) and *naganoi* species groups (Kons et al., 2017) because we were only able to sample a single representative from each. Additionally, among the 20 Groups defined by Barnes and McDunnough (1918), our analyses independently recovered 10 of those 15 Groups that could be tested; we also recovered their Groups V + IV as monophyletic, which they had only tentatively treated as separable. Lastly, we recovered the two foodplant groups proposed by Ishizuka et al. (2015: 159) as monophyletic, although their included species differed nominally from ours (13 Salicaceae-feeding species, corresponding to the *nupta* group sensu Borth, Kons, Saldaitis, & Gall, 2017; and 29 Juglandaceae-feeding species, corresponding to our node 17, Figure 2).

Topological hypothesis testing

The placement of *U. macula* as sister to *Catocala* was supported by 89.4% of FcLM quartets of the concatenated nucleotide dataset (Figure S7). Results of this analysis corroborate results from our phylogenetic analyses, with alternative topologies having limited support. Specifically, 6.4% of quartets pair *U. macula* with outgroups, and 1.7% pair *U. macula* with other *Ulotrichopus* species. Relationships were unresolved in 2.5% of quartets.

Evolution of larval host plant associations

Our ancestral state reconstruction resolved Fabaceae as the likely ancestral larval host of *Catocala* with a 96.3% probability, and *Catocala* + *Ulotrichopus*, with an 83.9% probability (Figure 2). The most common plant family in our dataset was Fagaceae, which is used by 28% of all the species for which larval host plant family could be determined (Table S2). Fagaceae-feeding arose at least four times on *Quercus* from Fabaceae feeding ancestors (nodes 4, 6, 11 and 15; Figure 2), and we observed a single shift from *Quercus* to *Fagus* (node 7).

Salicaceae-feeding arose three times, as observed in both the nucleotide (Figures 1, 2) and ASTRAL (Figure S6) analyses, and twice in the amino acid analysis (where Salicaceae-feeding formed a monophyletic group with low support: UFBoot/SH-aLRT = 83/67.8). Rosaceae-feeding arose three times, with Rosaceae-feeding clades A (node 12), and B (node 14), and *C. bella* Butler likely descending from a Fabaceae-feeding ancestor. Ericaceae-feeding (node 16) and Juglandaceae-feeding (node 17) appear to have arisen from a shared Fabaceae-feeding ancestor, and Myricaceae-feeding (node 18) may have arisen from an ancestor that fed on Juglandaceae. Our results suggest that the Betulaceae-feeding *C. ella* Butler arose from a *Juglandaceae*-feeding ancestor, and Malvaceae-feeding (node 10) and Ulmaceae-feeding (node 8) both likely arose from Fabaceae-feeding ancestors.

DISCUSSION

Our study provides the first robust phylogenetic framework for *Catocala* underwing moths and relatives. We recovered *Ulotrichopus* as the most closely related genus to *Catocala*, in agreement with the results of Mitter and Silverfine (1988), Kühne (2005) and Zahiri et al. (2012). However, we recover *U. macula* as the sister species of *Catocala*, which differs from Mitter and Silverfine (1988) and Ishizuka et al. (2015).

Our larval host plant ancestral state reconstruction analysis also supports prior hypotheses (e.g., Gall, 1991a; Ishizuka et al., 2011, 2015; Sargent, 1978) that host plants have played an important role in the evolution and diversification of underwing moths. Nine of ten larval host plant families utilized by *Catocala* are rosids, eight of which belong to three closely related orders: Fabales, Fagales and Rosales (APG IV, 2016). The only non-rosid family utilized by *Catocala* is Ericaceae, which belongs to the asterid order Ericales. The ancestral state reconstruction analysis indicates that the ancestral larval host plant of *Catocala* and *Catocala* + *Ulotrichopus* was Fabacaeae, with probabilities of 96.3 and 83.9% respectively (Figure 2). Although we could not sample all described *Catocala* and *Ulotrichopus* species for inclusion into our phylogeny, many unsampled species, especially those in *Ulotrichopus*, feed on Fabaceae (Kravchenko et al., 2010; Kroon, 1999), providing further support for our findings.

There were 14 instances of larval host plant preference shifts from Fabaceae to other plant families. The most frequent shift was from Fabaceae to Fagaceae, which, based on our results, occurred independently at least four times (nodes 4, 6, 11, 15; Figure 2). Reasons for the frequent shifts remain unknown, but we postulate that it was to exploit abundant food resources having similar general plant chemistry. Fabaceae and Fagaceae share phenolic acids and flavonoids (van der Linden et al., 2021). Fagaceae dominate many biomes from semiarid savannah to moist deciduous forests across the globe (e.g., Manos & Hipp, 2021; Narango et al., 2020; Stevens, 2017), and are larval host plants for a vast number of Lepidoptera species worldwide.

The earliest host plant shifts may have occurred in North Africa and the Mediterranean, from a Fabaceae-feeding *Ulotrichopus* inhabiting savannah-like habitats, that gave rise to Palearctic Fagaceae-feeding *Catocala* nodes 4 and 6 that use *Quercus* (with a subsequent shift to *Fagus* at node 7). All species in the diverse Nearctic Fagaceae-feeding clade that originated from a shift from Fabaceae at node 15 feed on *Quercus*.

Many species belonging to the Salicaceae-feeding clades (nodes 5, 13; Figure 2) occur in forested habitats across Eurasia and North America. The short internodes appearing in the clade of North American species sister to *C. unijuga* Walker, suggest that Salicaceae-feeding allowed this clade to take advantage of a range of unoccupied but suitable ecological niches in western North America such as desert riparian habitats. Willows (*Salix*) and poplars (*Populus*), the primary plant genera utilized by Salicaceae-feeding *Catocala*, are widespread in both temperate and boreal forests throughout the Holarctic (Brubaker et al., 2005). The shift to Salicaceae may have allowed expansion into higher latitudes, facilitating the colonisation of North America. Nodes 5 and 13 also include several Eurasian Salicaceae-feeding taxa associated with desert riparian and oasis habitats. Additional sampling of Eurasian *Catocala* is needed to assess whether similar rapid radiation may have occurred in such habitats.

The hyper-speciose Nearctic *Catocala* clade (nodes 14 through 19) includes 7 Fabaceae-feeding species that mostly inhabit savannah-like habitats, which are often found adjacent to habitats supporting Rosaceae, Juglandaceae and Fagaceae. In the midwestern United States, plums (*Malus*) often occur in savannahs, providing a food source for *Catocala clintonii* Grote, the sister group to the remainder of the Rosaceae-feeding clade originating at node 14 (this species uses *Crataegus* elsewhere in its geographic range). This Rosaceae-feeding clade and the Juglandaceae-feeding clade originating at node 19 are restricted to North America and comprise species that occur primarily in deciduous forests in the eastern part of the continent, suggesting that host plant shifts at both nodes occurred after the colonization of North America. Many of the short internodes in the Juglandaceae-feeding clade occur within a group of species specializing in hickories (*Carya*) (Gall, 1991b). Hickory-feeding species appear synchronically during the year (Sargent, 1976) and are often sympatric (ovipositing adult females of several species can be found on the same hickory plant Gall, 1991a, 1991b, 1991c; Schweitzer, 1982), suggesting that *Catocala* specialization on hickories is relatively recent. Such recent radiations onto different food plant groups in Nearctic *Catocala* are also consistent with the findings of Zahiri et al. (2017). In their comprehensive review of COI barcode patterns among 3,565 of the known 3,664 North American noctuoid species (747 total genera), they found that *Catocala* was a strikingly atypical noctuoid genus in having a disproportionate number (26 of 103) of species that share identical COI barcodes (an index of recent diversification). Nearly all (24 of 26) of the *Catocala* that share barcodes occur in two Nearctic-only speciose clades that we suggest are of recent origin, the *Carya* dominated Juglandaceae-feeders originating at node 19, and the Salicaceae-feeders originating at *Catocala marmorata* Edw. within node 5.

Species in the Ulmaceae-feeding clade (node 8) occur in diverse East Asian temperate forests where *Ulmus* and *Zelkova* occur (Fragnière et al., 2021). The shift to Ulmaceae-feeding from a Fabaceae-feeding ancestor likely occurred in East Asia due to the widespread availability of trees in this family, which subsequently may have facilitated diversification in the clade.

There were no shifts observed from other host plant families to Fabaceae, which could indicate that switching to another host plant family may cause host plant specialization that renders a backshift to Fabaceae less likely. This trend is also found in butterflies (Kawahara et al., 2023), which suggests that feeding on Fabaceae may be evolutionarily beneficial, preventing future back-shifts to other larval plant family hosts. Furthermore, there were multiple examples of host plant shifts at nodes that have short leading internode branch lengths. These are cases where all or nearly all species in these clades have retained an association with that plant family. A short internode branch length suggest that a single larval host plant shift led to rapid radiation, with the host plant shift possibly opening a new ecological niche leading to speciation.

In the Nearctic, considering all 747 noctuoid genera studied by Zahiri et al. (2017), *Catocala* stands out among the other speciose genera in having relatively limited larval food plant use. The number of Nearctic *Catocala* species (103) is exceeded only in *Euxoa* Hübner (182), *Sympistis* Hübner (177) and *Schinia* Hübner (126) and just four other genera contain more than 50 species in the Nearctic: *Acronicta* Ochs. (77), *Apamea* Ochs. (64), *Lacinipolia* McDunnough (61) and *Lithophane* Hübner (51). Nearctic *Catocala* larvae feed on a total of 20 host plant genera (Table S2). This contrasts with >40 host plant genera used in the Nearctic by each of *Euxoa*, *Schinia*, *Acronicta* and *Lithophane* (host data from moth species accounts in Handfield, 2011 and Wagner et al., 2011; larval biologies are less well documented for *Apamea*, *Sympistis* and *Lacinipolia*). On a per-moth-species basis, *Catocala* and *Schinia* appear to be larval host plant genus specialists, feeding on slightly >1 host plant genus per moth species, whereas *Euxoa* (>8) and *Acronicta* (>5) are larval host plant genus generalists. Considering each moth genus as a whole, *Catocala* has an average of >5 moth species per larval host plant genus, which is higher than that seen for *Schinia* (>1), *Euxoa* (<1) and *Acronicta* (<1). There are other salient host plant distinctions among these noctuoid genera (e.g., in the Nearctic all *Catocala* and most *Acronicta* feed on trees and woody shrubs whereas many *Schinia* and *Euxoa* feed on herbaceous plants), but it appears

that *Catocala* couples high taxonomic diversification with lower larval "host plant genus penetrance." The pattern is most pronounced (>10 *Catocala* species per larval host plant genus) in the Juglandaceae and Salicaceae feeding clades at nodes 5 and 19, which also contain the *Catocala* species that share identical COI barcodes (see above; Zahiri et al., 2017). We consider the Juglandaceae-feeding clade at node 19 to offer particular promise for exploring the relevance of this pattern to the evolution of host plant use.

Although we did not formally conduct a biogeographical analysis here, a few general trends appear from our findings. For example, the sister genus *Ulotrichopus* and all other genera in Catocalini are Afrotropical, suggesting a Palearctic origin of *Catocala*, possibly in North Africa prior to the aridification of the Sahara at least seven million years ago (Schuster et al., 2006). We recovered a strongly supported, higher-level clade of essentially exclusively Nearctic taxa (nodes 14 to 19) nested within *Catocala*, along with monophyletic groups of Nearctic species within otherwise Palearctic clades. This suggests that some *Catocala* might have reached North America via dispersal events, with one event occurring early in the evolutionary history of the genus, and several other events occurring more recently.

Lastly, the present study reconfirms that obtaining DNA from decades-old museum specimens is feasible (see e.g., Kawahara et al., 2023; Nunes et al., 2022). Half of our sequenced samples came from dry, pinned museum specimens and sequence capture was successful for 340 loci for a specimen of *C. abamita* Bremer & Grey collected at least 87 years prior to DNA extraction (Table S1). However, these sequences had a lower average sequence length per locus, compared to studies that utilized the same protocol but exclusively used alcohol-preserved samples (e.g., Homziak et al., 2019), suggesting that while sequence capture of old pinned museum specimens is possible, better results can be obtained from alcohol-preserved specimens. The Lep1 target capture kit has been successful at resolving relationships of moths that are less than ~21 Ma old (Johns et al., 2018), and in our ML tree, 12/168 (7%) of nodes had weak support (UFBoot <95 and SH-aLRT <80; Figure 1). Although we do not know what accounted for the weak support for these nodes, we believe this may be due to many of the LEP1 loci are short (<200 bp). Short loci will lead to poor support when using coalescent-based analyses, which is what we observed (Figures S5, S6). Future studies on *Catocala* phylogenetics should increase taxon sampling to provide a complete picture of the evolution of this charismatic insect lineage.

AUTHOR CONTRIBUTIONS

Nicholas T. Homziak: Conceptualization; data curation; formal analysis; visualization; writing – original draft; writing – review and editing; methodology; project administration; investigation. **Caroline G. Storer:** Supervision; data curation; formal analysis; writing – original draft; methodology; resources; writing – review and editing. **Lawrence F. Gall:** Conceptualization; data curation; visualization; writing – original draft; investigation; supervision; validation; writing – review and editing. **Robert J. Borth:** Validation; writing – review and editing; writing – original draft; visualization; data curation; conceptualization; supervision; investigation. **Akito Y. Kawahara:** Funding acquisition; resources; project administration; writing – review and editing; writing – original draft; visualization; formal analysis; data curation; conceptualization; methodology; investigation; supervision.

ACKNOWLEDGEMENTS

The authors thank Kelly Dexter, Amanda Markee, Rebeccah Messcher and David Plotkin for help with molecular lab work, specimen preparation and analysis. Benjamin Baiser, Phil Hahn, Jennifer Gillett-Kaufman, Charles Mitter and Emily Sessa helped contribute to the design of this study. Hugo Kons, Jr. assisted with specimen identifications and provided input that improved the quality of this manuscript. Marianne Espeland, Reza Zahiri and an anonymous referee provided valuable comments that improved the quality of the manuscript. The University of Florida Research Computing (http://researchcomputing.ufl.edu) provided computational resources and support for this research. An NSF GRFP award (DGE-1315138) to NTH supported this research at the University of Florida.

CONFLICT OF INTEREST STATEMENT

All authors certify that they have no affiliation with or involvement in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript.

DATA AVAILABILITY STATEMENT

Data supporting the findings of this study are openly available at NCBI (Accession PRJNA963183) and Dryad (doi:10.5061/dryad.n2z34tn1v).

ORCID

Nicholas T. Homziak https://orcid.org/0000-0002-8855-1303
Caroline G. Storer https://orcid.org/0000-0002-0349-0653
Lawrence F. Gall https://orcid.org/0000-0003-0690-0969
Robert J. Borth https://orcid.org/0000-0002-5228-5699
Akito Y. Kawahara https://orcid.org/0000-0002-3724-4610

REFERENCES

APG IV. (2016) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181, 1–20. Available from: https://doi.org/10.1111/boj.12385

Barnes, W. & McDunnough, J. (1918) Illustrations of the North American species of the genus *Catocala*. *Memoirs of the American Museum of Natural History*, n.s. III, 3–47.

Borth, R.J. & Kons, H.L., Jr. (2016) A new species of *Catocala* (Lepidoptera: Noctuidae) from California. *Bulletin of the Peabody Museum of Natural History*, 57, 241–252. Available from: https://doi.org/10.3374/014.057.0201

Borth, R.J., Kons, H.L., Jr. & Saldaitis, A. (2017) A new species of *Catocala* (Lepidoptera: Noctuidae) from Vietnam. *Bulletin of the Peabody Museum of Natural History*, 58, 47–64. Available from: https://doi.org/10.3374/014.058.0104

Borth, R.J., Kons, H.L., Jr., Saldaitis, A. & Gall, L.F. (2017) The taxonomy of *Catocala nupta* (Linnaeus, 1767) and its allies, with description of a new species (Lepidoptera: Noctuidae). *Zootaxa*, 4263, 251–272. Available from: https://doi.org/10.11646/zootaxa.4263.2.3

Breinholt, J.W., Earl, C., Lemmon, A.R., Lemmon, E.M., Xiao, L. & Kawahara, A.Y. (2018) Resolving relationships among the

megadiverse butterflies and moths with a novel pipeline for anchored phylogenomics. *Systematic Biology*, 67, 78–93.

Brubaker, L.B., Anderson, P.M., Edwards, M.E. & Lozhkin, A.V. (2005) Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data. *Journal of Biogeography*, 32, 848–883. Available from: https://doi.org/10.1111/j.1365-2699.2004.01203.x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST plus: architecture and applications. *BMC Bioinformatics*, 10, 421. Available from: https://doi.org/10.1186/1471-2105-10-421

Chang, Z., Li, G.J., Liu, J.T., Zhang, Y., Ashby, C., Liu, D.L. et al. (2015) Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biology*, 16, 30. Available from: https://doi.org/10.1186/S13059-015-0596-2

Cho, S., Epstein, S.W., Mitter, K.T., Hamilton, C.A., Plotkin, D., Mitter, C. et al. (2016) Preserving and vouchering butterflies and moths for large-scale museum-based molecular research. *PeerJ*, 4, e2160. Available from: https://doi.org/10.7717/peerj.2160

Dubatolov, V. & Kosterin, O. (2000) Nemoral species of Lepidoptera (Insecta) in Siberia: a novel view on their history and the timing of their range disjunctions. *Entomologica Fennica*, 11, 141–166. Available from: https://doi.org/10.33338/ef.84061

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461.

Fibiger, M. & Hacker, H.H. (2005) Systematic list of the Noctuidae of Europe (Notodontidae, Nolidae, Arctiidae, Lymantriidae, Erebidae, Micronoctuidae, and Noctuidae). *Esperiana*, 11, 93–205. Available from: http://esperiana.net/mediapool/86/862516/data/Esperiana_Band_11_93-206.pdf

Fragnière, Y., Song, Y.G., Fazan, L., Manchester, S.R., Garfì, G. & Kozlowski, G. (2021) Biogeographic overview of Ulmaceae: diversity, distribution, ecological preferences, and conservation status. *Plants (Basel, Switzerland)*, 10(6), 1111. Available from: https://doi.org/10.3390/plants10061111

Gall, L.F. (1991a) Evolutionary ecology of sympatric *Catocala* moths (Lepidoptera: Noctuidae) III. Experiments on female oviposition preference. *Journal of Research on the Lepidoptera*, 29, 217–233. Available from: https://www.biodiversitylibrary.org/page/53314536

Gall, L.F. (1991b) Evolutionary ecology of sympatric *Catocala* moths (Lepidoptera: Noctuidae) I. experiments on larval foodplant specificity. *Journal of Research on the Lepidoptera*, 29, 173–194. Available from: https://www.biodiversitylibrary.org/page/53314492

Gall, L.F. (1991c) Evolutionary ecology of sympatric *Catocala* moths (Lepidoptera: Noctuidae) II. Sampling for wild larvae on their foodplants. *Journal of Research on the Lepidoptera*, 29, 195–216. Available from: https://www.biodiversitylibrary.org/page/53314514

Gall, L.F. & Hawks, D.C. (2010) Systematics of moths in the genus *Catocala* (Lepidoptera, Erebidae) IV. Nomenclatorial stabilization of the Nearctic fauna, with a revised synonymic check list. *Zookeys*, 39, 37–83. Available from: https://doi.org/10.3897/zookeys.39.425

Goater, B., Ronkay, L. & Fibiger, M. (2003) *Noctuidae Europaeae 10: Catocalinae and Plusiinae*. Soro: Entomological Press.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307–321. Available from: https://doi.org/10.1093/sysbio/syq010

Hamilton, C.A., St Laurent, R.A., Dexter, K., Kitching, I.J., Breinholt, J.W., Zwick, A. et al. (2019) Phylogenomics resolves major relationships and reveals significant diversification rate shifts in the evolution of silk moths and relatives. *BMC Evolutionary Biology*, 19, 182. Available from: https://doi.org/10.1186/s12862-019-1505-1

Hampson, G.F. (1913) *Catalogue of the Lepidoptera Phalaenae of the British museum*, Vol. 12. London: Trustees of the British Museum. Available from: https://www.biodiversitylibrary.org/item/101378

Handfield, L. (2011) *Le Guide des Papillons du Quebec. Volume 1, Part 1*. Broucherville: Broquet, Inc.

Hardwick, D.F. (1950) Preparation of slide mounts of Lepidopterous genitalia. *The Canadian Entomologist*, 82, 231–235. Available from: https://doi.org/10.4039/Ent82231-11

Hawks, D.C. (2010) Review of the *Catocala delilah* species complex (Lepidoptera, Erebidae). *ZooKeys*, 39, 13–35. Available from: https://doi.org/10.3897/zookeys.39.439

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. & Vinh, L.S. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.

Holland, W.J. (1903) *The moth book; a popular guide to a knowledge of the moths of North America*. New York: Doubleday, Page & Co. Available from: https://www.biodiversitylibrary.org/item/217057

Homziak, N.T., Breinholt, J.W., Branham, M.A., Storer, C.G. & Kawahara, A.Y. (2019) Anchored hybrid enrichment phylogenomics resolves the backbone of erebine moths. *Molecular Phylogenetics & Evolution*, 131, 99–105. Available from: https://doi.org/10.1016/j.ympev.2018.10.038

Hsu, Y.-F., Wang, L.-H., Huang, C.-L., Braby, M.F., Lin, W.-J., Lin, R.-J. et al. (2021) Discovery of specialist beech-feeding underwing moths from Taiwan (Lepidoptera: Erebidae: Catocalinae) and its implication to biodiversity conservation of Taiwan beech forests. *Annals of the Entomological Society of America*, 114, 338–354. Available from: https://doi.org/10.1093/aesa/saaa062

Hundsdoerfer, A.K. & Kitching, I.J. (2010) A method for improving DNA yield from century-plus old specimens of large lepidoptera while minimizing damage to external and internal abdominal characters. *Arthropod Systematics & Phylogeny*, 68, 151–155.

Ishizuka, K. (2011) *Catocala* of the World. Handbook Series of Insects 1. Mushi-Sha, Tokyo.

Ishizuka, K., Matsuo, M. & Nonaka, M. (2015) Molecular phylogenetic analysis of *Catocala* moths based on the nuclear ITS2 and 28S rRNA gene sequences (Lepidoptera, Noctuidae). *Tinea*, 23, 157–170. Available from: https://ci.nii.ac.jp/naid/40020692047/en/

Ishizuka, K., Shinkawa, T., Endo, H. & Nonaka, M. (2011) Molecular phylogeny of Japanese *Catocala* moths based on nucleotide sequences of the mitochondrial ND5 gene. *Zoological Science*, 28, 910–915. Available from: https://doi.org/10.2108/zsj.28.910

Johns, C.A., Toussaint, E.F.A., Breinholt, J.W. & Kawahara, A.Y. (2018) Origin and macroevolution of micro-moths on sunken Hawaiian islands. *Proceedings of the Royal Society B: Biological Sciences*, 285, 20181047. Available from: https://doi.org/10.1098/rspb.2018.1047

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. Available from: http://mbe.oxfordjournals.org/content/30/4/772.full.pdf

Kawahara, A.Y., Plotkin, D., Espeland, M., Meusemann, K., Toussaint, E.F.A., Donath, A. et al. (2019) Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proceedings of the National Academy of Sciences, USA*, 116, 22657–22663. Available from: https://doi.org/10.1073/pnas.1907847116

Kawahara, A.Y., Storer, C., Carvalho, A.P.S., Plotkin, D.M., Condamine, F.L., Braga, M.P. et al. (2023) A global phylogeny of butterflies reveals their evolutionary history, ancestral hosts and biogeographic origins. *Nature Ecology and Evolution*. Available from: https://doi.org/10.1038/s41559-023-02041-9

Kononenko, V. (2010) *Noctuidae Sibiricae. Volume 2. Micronoctuidae, Noctuidae: Rivulinae – Agarsitinae*. Soro: Entomological Press.

Kons, H.L., Jr. & Borth, R.J. (2015a) A new species of *Catocala* (Lepidoptera: Noctuidae) from Florida. *Bulletin of the Peabody Museum of Natural History*, 56, 67–79. Available from: https://doi.org/10.3374/014.056.0104

Kons, H.L., Jr. & Borth, R.J. (2015b) A new species of *Catocala* (Lepidoptera: Noctuidae) from the southeastern United States. *Bulletin of the Peabody Museum of Natural History*, 56, 55–65. Available from: https://doi.org/10.3374/014.056.0103

Kons, H.L., Jr. & Borth, R.J. (2016) A new species of *Catocala* (Lepidoptera: Noctuidae) from the Gulf coast of the Florida panhandle. *Bulletin of the Peabody Museum of Natural History*, 57, 229–239. Available from: https://doi.org/10.3374/014.057.0207

Kons, H.L., Jr. & Borth, R.J. (2017) A new species of *Catocala* (Lepidoptera: Noctuidae) from the south Central United States. *Bulletin of the Peabody Museum of Natural History*, 58, 31–46. Available from: https://doi.org/10.3374/014.058.0101

Kons, H.L., Jr., Borth, R.J. & Saldaitis, A. (2016) A new species of *Catocala* Schrank, 1802 (Lepidoptera: Noctuidae) from Primorsky Krai, Russia. *Zootaxa*, 4105, 389–399. Available from: https://doi.org/10.11646/zootaxa.4105.4.7

Kons, H.L., Jr., Borth, R.J., Saldaitis, A. & Didenko, S. (2017) The *Catocala naganoi* species group (Lepidoptera: Noctuidae), with a new species from Vietnam. *Zootaxa*, 4358, 79–106. Available from: https://doi.org/10.11646/zootaxa.4358.1.3

Kravchenko, V., Matov, A., Speidel, W., Rybalov, L., Seyoum, E., Beredin, S. et al. (2010) Contributions to the Lepidoptera fauna of Ethiopia: Catocalinae (Noctuidae). *Lepidoptera Novae (Gainesville)*, 3, 41–52.

Kroon, D.M. (1999) Lepidoptera of southern Africa. In: *Host-plants and other associations: a catalogue. Lepidopterists*. Jukskei Park, South Africa: Society of Southern Africa.

Kück, P. & Longo, G.C. (2014) FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology*, 11, 81. Available from: https://doi.org/10.1186/s12983-014-0081-x

Kühne, L. (2005) Revision und Phylogenie der Gattungsgruppe Crypsotidia Rothschild, 1901, Tachosa Walker, 1869, Hypotacha Hampson, 1913, Audea Walker, [1858] 1857, und Ulotrichopus Wallengren, 1860 (Lepidoptera, Noctuidae, Catocalinae). *Esperiana Memoir*, 2, 7–220.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, 14, 82. Available from: https://doi.org/10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773. Available from: https://doi.org/10.1093/molbev/msw260

Lees, D. & Zilli, A. (2019) *Moths – a complete guide to biology and behavior*. Washington, DC: Smithsonian Books.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61, 727–744. Available from: https://doi.org/10.1093/sysbio/sys049

Manos, P. & Hipp, A.L. (2021) An updated infrageneric classification of the north American oaks (*Quercus* subgenus *Quercus*): review of the contribution of phylogenomic data to biogeography and species diversity. *Forests*, 12, 786. Available from: https://doi.org/10.3390/f12060786

Minh, B.Q., Nguyen, M.A.T. & von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30, 1188–1195. Available from: https://doi.org/10.1093/molbev/mst024

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534. Available from: https://doi.org/10.1093/molbev/msaa015

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30, I541–I548. Available from: https://doi.org/10.1093/bioinformatics/btu462

Mitter, C. & Silverfine, E. (1988) On the systematic position of *Catocala* Schrank (Lepidoptera: Noctuidae). *Systematic Entomology*, 13, 67–84. Available from: https://doi.org/10.1111/j.1365-3113.1988.tb00230.x

Narango, D.L., Tallamy, D.W. & Shropshire, K.J. (2020) Few keystone plant genera support the majority of Lepidoptera species. *Nature Communications*, 11, 5751. Available from: https://doi.org/10.1038/s41467-020-19565-4

Nunes, R., Storer, C., Doleck, T., Kawahara, A.Y., Pierce, N.E. & Lohman, D.J. (2022) Predictors of sequence capture in a large-scale anchored phylogenomics project. *Frontiers in Ecology & Evolution*, 10, 943361. Available from: https://doi.org/10.3389/fevo.2022.943361

Revell, L.J. (2012) Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology & Evolution*, 3, 217–223. Available from: https://doi.org/10.1111/j.2041-210X.2011.00169.x

Sargent, T.D. (1976) *Legion of night: the underwing moths*. Amherst: University of Massachusetts Press.

Sargent, T.D. (1978) On the maintenance of stability in hindwing diversity among moths of the genus *Catocala* (Lepidoptera: Noctuidae). *Evolution*, 32, 424–434. Available from: https://doi.org/10.2307/2407609

Schuster, M., Duringer, P., Ghienne, J.-F., Vignaud, P., Mackaye, H.T., Likius, A. et al. (2006) The age of the Sahara desert. *Science*, 311, 821 LP – 821. Available from: https://doi.org/10.1126/science.1120161

Schweitzer, D.F. (1982) Field observations of foodplant overlap among sympatric *Catocala* feeding on Juglandaceae. *Journal of the Lepidopterists' Society*, 36, 256–263. Available from: https://www.biodiversitylibrary.org/item/128080

Sohn, J.C., Labandeira, C.C. & Davis, D.R. (2015) The fossil record and taphonomy of butterflies and moths (Insecta, Lepidoptera): implications for evolutionary diversity and divergence-time estimates. *BMC Evolutionary Biology*, 15, 1–15. Available from: https://doi.org/10.1186/s12862-015-0290-8

St Laurent, R.A., Hamilton, C.A. & Kawahara, A.Y. (2018) Museum specimens provide phylogenomic data to resolve relationships of sackbearer moths (Lepidoptera, Mimallonoidea, Mimallonidae). *Systematic Entomology*, 43, 729–761. Available from: https://doi.org/10.1111/syen.12301

Staude, H.S., Maclean, M., Mecenero, S., Pretorius, R.J., Oberprieler, R.G., Van Noort, S. et al. (2020) An overview of Lepidoptera-host-parasitoid associations for southern Africa, including an illustrated report on 2370 African Lepidoptera-host and 119 parasitoid-Lepidoptera associations. *Metamorphosis*, 31, 1–380.

Staude, H., Mecenero, S., Oberprieler, R., Sharp, A., Sharp, I., Williams, M. et al. (2016) An illustrated report on the larvae and adults of 962 African Lepidoptera species. Results of the Caterpillar rearing group: a novel, collaborative method of rearing and recording lepidopteran life-histories: Noctuoidea: Erebidae: others. *Metamorphosis*, 27, 165–188.

Stevens, P.F. (2017) Angiosperm Phylogeny Website. Version 14, July 2017 [with subsequent updates]. http://www.mobot.org/MOBOT/research/APweb/ Accessed 2 December 2022.

Strimmer, K. & von Haeseler, A. (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences, USA*, 94, 6815–6819.

van der Linden, C.F.H., Wallis, M.F. & DeVries, S.S. (2021) Great chemistry between us: the link between plant chemical defenses and butterfly evolution. *Ecology & Evolution*, 11, 8595–8613. Available from: https://doi.org/10.1002/ece3.7673

van Nieukerken, E.J., Kaila, L., Kitching, I.J., Kristensen, N.P., Lees, D.C., Minet, J. et al. (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang, Z.-Q. (Ed.), Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. *Zootaxa*, 3148, 212–221.

Wagner, D.L., Schweitzer, D.F., Sullivan, J.B. & Reardon, R.C. (2011) *Owlet caterpillars of eastern North America (Lepidoptera: Noctuidae)*. Princeton: Princeton University Press.

Xia, Q.Y., Zhou, Z.Y., Lu, C., Cheng, D.J., Dai, F.Y., Li, B. et al. (2004) A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science*, 306, 1937–1940. Available from: https://doi.org/10.1126/science.1102210

Yang, Z., Plotkin, D., Landry, J.-F., Storer, C. & Kawahara, A.Y. (2021) Revisiting the evolution of Ostrinia moths with phylogenomics (Pyraloidea: Crambidae: Pyraustinae). *Systematic Entomology*, 46, 827–838. Available from: https://doi.org/10.1111/syen.12491

Zahiri, R., Holloway, J.D., Kitching, I.J., Lafontaine, J.D., Mutanen, M. & Wahlberg, N. (2012) Molecular phylogenetics of Erebidae (Lepidoptera, Noctuoidea). *Systematic Entomology*, 37, 102–124. Available from: https://doi.org/10.1111/j.1365-3113.2011.00607.x

Zahiri, R., Lafontaine, J.D., Schmidt, B.C., deWaard, J.R., Zakharov, E.V. & Hebert, P.D.N. (2017) Probing planetary biodiversity with DNA barcodes: the Noctuoidea of North America. *PLoS One*, 12(6), e0178548. Available from: https://doi.org/10.1371/journal.pone.0178548

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19, 153. Available from: https://doi.org/10.1186/s12859-018-2129-y

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Subset of best scoring maximum likelihood tree showing relationships between *Catocala*, *Ulotrichopus* and outgroups inferred from 685 loci from concatenated nucleotide dataset. Node support values were calculated using both the SH-like approximate likelihood ratio test (SH-aLRT) and ultrafast bootstrap approximation (UFBoot). Branch support is shown as SH-aLRT/UFBoot.

Figure S2. Subset of best scoring maximum likelihood tree showing relationships within *Catocala* and *Ulotrichopus* inferred from 685 loci from concatenated nucleotide dataset. Node support values were calculated using both the SH-like approximate likelihood ratio test (SH-aLRT) and ultrafast bootstrap approximation (UFBoot). Branch support is shown as SH-aLRT/UFBoot.

Figure S3. Subset of best scoring maximum likelihood tree showing relationships between *Catocala*, *Ulotrichopus* and outgroups inferred from 685 loci from the concatenated amino acid dataset. Node support values were calculated using both the SH-like approximate likelihood ratio test (SH-aLRT) and ultrafast bootstrap approximation (UFBoot). Branch support is shown as SH-aLRT/UFBoot.

Figure S4. Subset of best scoring maximum likelihood tree showing relationships within *Catocala* and *Ulotrichopus* inferred from 685 loci from the concatenated amino acid dataset. Node support values were calculated using both the SH-like approximate likelihood ratio test (SH-aLRT) and ultrafast bootstrap approximation (UFBoot). Branch support is shown as SH-aLRT/UFBoot.

Figure S5. Subset of species tree showing relationships between *Catocala*, *Ulotrichopus* and outgroups inferred from gene trees of 685 loci using ASTRAL. Local posterior probability (LPP) values are shown at each node.

Figure S6. Subset of species tree showing relationships within *Catocala* and *Ulotrichopus* inferred from gene trees of 685 loci using ASTRAL. Posterior probability values are shown at each node.

Figure S7. Output of four cluster likelihood mapping analyses implemented in IQ-TREE to test the placement of *Ulotrichopus macula* relative to *Catocala* and other *Ulotrichopus* species in our study. Numbers indicate the following clusters of taxa: (1) *U. macula*, (2) *Catocala* sensu stricto, (3) African *Ulotrichopus* and (4) Outgroups. Numerical values around the upper triangle represent the three alternative quartet topologies being tested. Percent support values for each quartet are shown at the corners of each lower triangle. Support values in the middle and along the sides of the lower right triangle represent percentages of unresolved quartets.

Figure S8. Node numbers shown on the phylogeny are used as input for the ancestral state reconstruction of larval host plant preference in SIMMAP. Probabilities associated with each node number are shown in Table S3.

Table S1. List of taxa sampled in this study, including voucher numbers.

Table S2. Larval hostplant dataset used for ancestral state reconstruction analysis.

Table S3. Results of SIMMAP analysis showing the probabilities of larval host plant family ancestral states at each node.

How to cite this article: Homziak, N.T., Storer, C.G., Gall, L.F., Borth, R.J. & Kawahara, A.Y. (2023) Phylogenomics resolves major relationships of *Catocala* underwing moths. *Systematic Entomology*, 1–11. Available from: https://doi.org/10.1111/syen.12595