Biology Faculty Publications and Presentations | Department of Biological Sciences

9-18-2019

Phylogenomics Resolves Major Relationships and Reveals Significant Diversification Rate Shifts in the Evolution of Silk Moths and Relatives

J. R. Barber
Boise State University

For a complete list of authors, please see article.

Phylogenomics resolves major relationships and reveals significant diversification rate shifts in the evolution of silk moths and relatives

C. A. Hamilton[1,2*] , R. A. St Laurent[1] , K. Dexter[1] , I. J. Kitching[3] , J. W. Breinholt[1,4] , A. Zwick[5] , M. J. T. N. Timmermans[6] , J. R. Barber[7†] and A. Y. Kawahara[1*†]

Abstract

Background: Silkmoths and their relatives constitute the ecologically and taxonomically diverse superfamily Bombycoidea, which includes some of the most charismatic species of Lepidoptera. Despite displaying spectacular forms and diverse ecological traits, relatively little attention has been given to understanding their evolution and drivers of their diversity. To begin to address this problem, we created a new Bombycoidea-specific Anchored Hybrid Enrichment (AHE) probe set and sampled up to 571 loci for 117 taxa across all major lineages of the Bombycoidea, with a newly developed DNA extraction protocol that allows Lepidoptera specimens to be readily sequenced from pinned natural history collections.

Results: The well-supported tree was overall consistent with prior morphological and molecular studies, although some taxa were misplaced. The bombycid Arotros Schaus was formally transferred to Apatelodidae. We identified important evolutionary patterns (e.g., morphology, biogeography, and differences in speciation and extinction), and our analysis of diversification rates highlights the stark increases that exist within the Sphingidae (hawkmoths) and Saturniidae (wild silkmoths).

Conclusions: Our study establishes a backbone for future evolutionary, comparative, and taxonomic studies of Bombycoidea. We postulate that the rate shifts identified are due to the well-documented bat-moth "arms race". Our research highlights the flexibility of AHE to generate genomic data from a wide range of museum specimens, both age and preservation method, and will allow researchers to tap into the wealth of biological data residing in natural history collections around the globe.

Keywords: Anchored Hybrid Enrichment, Targeted sequence capture, Phylogenomics, Bombycoidea, Lepidoptera, Natural history collections

Background

The Bombycoidea include some of the most charismatic moths among all Lepidoptera. This ecologically diverse superfamily comprises ten families, 520 genera, and 6092 species [1]. Although widespread globally, the highest diversity of bombycoids occurs in the tropics and subtropics. This diversity includes the most spectacular forms

(i.e., range of body sizes and wing shapes) and functions (i.e., mimicry, predator avoidance, flight capabilities, and feeding strategies) in the Lepidoptera [2].

A number of bombycoids have become important contributors to human culture, originally as economically important species for sericulture or as agricultural pests, but more recently as model organisms for comparative studies of genetics, development, and physiology [2]. Additionally, many lineages play important roles as pollinators ([3–9]), or as indicators in biodiversity and habitat quality assessments [10]. Of the 10 families, three contain species that have

* Correspondence: hamiltonlab@uidaho.edu; kawahara@flmnh.ufl.edu
†J. R. Barber and A. Y. Kawahara are senior authors.
[1]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
Full list of author information is available at the end of the article

been used as model organisms (Bombycidae, Saturniidae, and Sphingidae). Unfortunately, a strong understanding of bombycoid familial relationships has remained largely elusive. For example, Bombycidae have been considered the sister lineage to either Saturniidae or Sphingidae, depending on data types and methodological approaches ([11–18]); but see Breinholt and Kawahara [19], and few studies have included representatives from all the major bombycoid lineages.

Despite their charisma and intrigue, the lack of a robust phylogeny based on broad and dense taxon sampling across the Bombycoidea is dramatically affecting our ability to answer fundamental questions about the evolutionary drivers of their diversity, in particular, species number, body size, ecological niche, etc. Monophyly of the Bombycoidea has been supported by six morphological synapomorphies [15], but Zwick [20] determined that only two of these were systematically informative: one poorly understood thoracic character [21], and one relating to the arrangement of forewing veins (A. Zwick, unpublished). Recent molecular studies of bombycoid systematics ([11, 13, 20, 22]) have resulted in substantial differences in terms of relationships from morphology-based phylogenetic hypotheses ([15, 21, 23]). To date, nearly all molecular studies of bombycoids have included fewer than 20 protein-coding genes for ≤50 species (e.g., [11, 13, 22, 24, 25]). Although these studies agreed on the monophyly of the superfamily, many relationships among families, subfamilies, and tribes remained unclear and were characterized by weak branch support or conflicting signal. Modern phylogenomics (e.g., based on Anchored Hybrid Enrichment) has been shown to be effective to resolve relationships among Lepidoptera at multiple taxonomic levels ([16, 26–29]) including the Bombycoidea ([16, 30]). However, those studies that utilize phylogenomics to resolve inter- and intra-familial relationships within the Bombycoidea have been limited by taxon sampling.

To better answer questions regarding bombycoid evolution, we applied Anchored Hybrid Enrichment (AHE) targeted-sequencing phylogenomics [31]. We developed a new, bombycoid-specific, AHE probe set (here called "BOM1"), redesigned from the "LEP1" probe set of Breinholt et al. [16]. Our new probe set captures "legacy" Sanger sequencing-based loci that were part of existing bombycoid molecular datasets ([11, 13, 25, 32]), enabling the merging of older published datasets with those generated from the BOM1 probe set. To further improve the ability to generate a dataset with greater taxon and locus sampling, we developed a new method for extracting DNA from pinned natural history specimens and show that our extraction approach is successful in obtaining DNA sequence data for phylogenomics. In total, our dataset resulted in 571 loci for 117 species across Bombycoidea.

We also use this opportunity to utilize the phylogeny to examine patterns of diversification in the superfamily. Bombycoidea are well-known to have multiple different ecological life-history strategies, especially for the Saturniidae and Sphingidae – two lineages that harbor the majority of described bombycoid species ([30, 33–35]). The divergent life-history strategies of these two families ([36, 37]) has likely played a major role in driving their diversity. For example, the majority of hawkmoths feed as adults, seeking out nectar resources during their relatively long lives (weeks to months). During this time, females experience multiple mating events and retain the eggs internally for long periods to allow egg maturation and host plant discovery [36]. This ecological strategy is significantly different from saturniids ([19, 24, 38]), which depend entirely upon the resources acquired during the larval period. Adult saturniids possess reduced or non-functional mouthparts and lay eggs almost immediately after mating. Furthermore, these lineages possess a number of different traits that appear to be anti-bat adaptations in response to echolocating bats, a lineage thought to have arisen approximately 60 million years ago ([39–43]). Hawkmoths, especially those in the Sphinginae and Macroglossinae, are strong fliers thought to have evolved hearing organs and ultrasound producing organs capable of jamming bat sonar ([34, 44]). Saturniids, in contrast, lack hearing organs and exhibit erratic evasive flight [36] and hindwing tails that deflect bat echoes ([30, 35]). These traits firmly establish the relative ecological roles of Saturniidae and Sphingidae as an important natural experiment from which we can gather valuable information regarding the evolution of predator/prey interactions in moths and their ultimate effects on the diversification process in general.

Methods

The "BOM1" AHE probe set design

Anchored Hybrid Enrichment (AHE) is a targeted-sequencing methodology designed to capture hundreds of unique orthologous loci (i.e., single copy, phylogenetically-informative markers) from across the genome, for resolving both shallow and deep-level evolutionary relationships ([11, 13]). Probes are designed to anchor in conserved regions that are flanked by variable regions randomly spread throughout the genome. This approach creates a diverse set of informative loci that include exons, introns, intergenic, and conserved regions of the genome. Targeted-sequencing approaches, like AHE and UCE (Ultraconserved Elements), provide mechanisms whereby different researchers can confidently and effectively use the same loci for independent projects, allowing for the combination of data across studies.

Breinholt et al. [16] constructed a Lepidoptera Agilent Custom SureSelect Target Enrichment "LEP1" probe kit, designed for 855 loci. However, this probe set is not

specific to Bombycoidea, and does not include some of the traditional loci that have been used to study the phylogenetics of Bombycoidea. In order to build a more Bombycoidea-specific AHE probe set and phylogenomic dataset, we began by modifying the LEP1 kit, evaluating which loci were most phylogenetically-informative within the superfamily, optimizing the set of probes to recover these loci, and including 24 previously-sequenced Sanger-sequenced loci, e.g., CO1, CAD, DDC, period, wingless and others from [11, 13, 25, 32], as well as eight vision-related genes (see Additional file 1: Table S1 for locus names). The probes for these vision-related genes are based on phototransduction genes, mined from eye or head transcriptomes (unpublished; generated by AYK), and are included for future analyses to investigate their evolution across the superfamily.

To determine the informative loci for BOM1, phylogenetically-informative loci were identified by examining the sequence variation of 56 bombycoid species from across the taxonomic breadth of the superfamily (55 AHE samples, from either Breinholt et al. [16] or generated for this study, plus loci mined from the *Bombyx mori* reference genome [45]). These samples were sequenced and processed using the LEP1 probe kit and Breinholt et al. [16] bioinformatics pipeline [46]. Individual gene trees were generated for each of the 855 loci, using Maximum Likelihood in RAxML v8.2 [47], under a GTRGAMMA model of evolution with 100 non-parametric bootstraps for node support. For each tree, phylogenetic informativeness was calculated using PhyDesign ([48–50]). Parsimony informative characters and the number of segregating sites were calculated using the R packages 'phyloch' [51] and 'ape' [52] respectively. Additionally, each phylogeny was scrutinized visually to determine whether branch lengths and topological patterns appeared realistic (i.e., no significant outliers present). Loci deemed to be phylogenetically uninformative, or those that were capturing poorly (< 60% of sampled species represented for a locus), were excluded from the probe set. The final BOM1 probe kit comprised 571 loci, 539 of which came from the original LEP1 kit.

Taxon sampling

To build a backbone phylogeny of the Bombycoidea, we sampled all major lineages (i.e., families, subfamilies, and tribes), with the exception of three rare, species-poor groups: subfamily Munychryiinae (Anthelidae) and tribes Sataspedini and Monardini (Sphingidae), for which representative samples were unavailable for DNA sequencing. Sampled lineages were chosen because: 1) they were appropriate representatives of the taxonomic group needed for the analysis (i.e., good morphological and evolutionary representative of a tribe); and 2) they were accessible for use in phylogenomics. In total, 115 ingroup

Bombycoidea species from 97 genera were included in the phylogenetic analysis, as well as two Lasiocampidae outgroups – the sister lineage to the bombycoids (see [38, 53]; Additional file 2: Table S2).

Specimens were obtained from fieldwork, historically preserved dry collections (Additional file 1: Table S1), and molecular tissue collections. Field-collected specimens were stored in ≥95% ethanol, RNAlater, or papered and dried with silica gel. Genomic DNA was extracted using Omni-Prep Genomic DNA Extraction Kits (G-Biosciences, St. Louis, MO, USA) and DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA, USA). DNA concentration was evaluated through agarose gel electrophoresis and fluorometry using a Qubit 2.0 (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA). Library preparation, hybridization enrichment, and Illumina HiSeq 2500 sequencing (PE100) was carried out at RAPiD Genomics (Gainesville, FL, USA). Specimen wing vouchering and tissue storage methods follow Cho et al. [54]. All DNA extracts and specimens preserved in ethanol, RNAlater, or those freshly papered were stored at – 80 °C at the Florida Museum of Natural History, McGuire Center of Lepidoptera and Biodiversity (MGCL).

DNA extraction protocol for museum specimens

We evaluated the efficacy of obtaining DNA from historical museum specimens because there is great interest in understanding the feasibility of this approach for use in phylogenetics and systematics ([28, 55–57]). The Lepidoptera specimens evaluated herein, were "field-pinned" (never rehydrated), "papered" (stored in an envelope and kept dry since collected, thus not rehydrated or pinned), and "traditionally-pinned" specimens (dried, rehydrated, and subsequently pinned) – the historically most common method of Lepidoptera specimen storage. Collecting dates ranged from 1987 to 2017 (Additional file 2: Table S2). Samples were not initially intended to be preserved for molecular sequencing and information about potential contaminants and/or extraction inhibitors (i.e. fumigation compounds, other chemicals) was unavailable. In some cases there was little soft-tissue to extract from within the abdomens, having degraded over time. Being that these samples were not kept in molecular-grade conditions, many were contaminated with fungal and bacterial growth, which could be identified visually by the spores left on the bodies or by the smell of decay. These factors along with the amount of sclerotized tissue present and the fact that the abdomens used needed to stay intact (not homogenized) for dissecting purposes, made extracting good quality genomic DNA challenging.

Our extraction method, detailed in the Additional file 10, attempts to account for several factors: the amount of degraded tissue, the presence of eggs, the relative fat content, and the overall abdomen size. Many commercial DNA extraction kits on the market (including the Omni

Prep kit used in this study) recommend using 10 mg–20 mg of well-preserved tissue for the extraction process. Given that the museum specimens used had been desiccated for many years, a number of abdomens had little to no visible internal soft tissue remaining. To digest the remaining material in solution, we increased the ratio of proteinase K to lysis buffer. Companies that produce DNA extractions kits know, as can be seen by their specific extraction kits, the relative fat content of samples is an issue because lipid-rich tissue can interfere with the digestion of the soft tissue as well as change the chemistry of the DNA isolation buffers. Specimens that appeared to be "greasy" or seemed to have an oily film on the abdomen were not used for extraction. The overall size of the abdomen was used to estimate the amount of lysis buffer needed in order to sufficiently submerge the abdomen and reach the available soft tissue. The amount of buffer also reflected the volume of reagents needed for the remainder of the DNA isolation process. Lastly, the modified protocol also allows a user to easily prepare the genitalia for taxonomic or morphological taxonomic work as these structures remain undamaged.

Bioinformatics

The bioinformatics pipeline of Breinholt et al. [16] was used to clean and assemble raw Illumina reads for each AHE locus. The pipeline uses a probe-baited iterative assembly that extends beyond the probe region, checks for quality and cross contamination due to barcode leakage, removes paralogs, and returns a set of aligned orthologs for each locus and taxon of interest. To accomplish these tasks, the pipeline uses the *Bombyx mori* genome [45], and an AHE reference library, which in this study was the BOM1 reference library.

Loci for phylogenetic analysis were selected by applying a cutoff of ≥40% sampled taxa recovery (i.e., for a locus to be included in the analysis, the locus had to be recovered in at least 40% of the sampled taxa). The pipeline evaluates density and entropy at each site of a nucleotide sequence alignment. We elected to trim with entropy and density cutoffs only in "flanking" regions, allowing the "probe" region (exon) to be converted into amino acid sequences. For a site (outside of the probe region) to remain, that site must pass a 60% density and 1.5 entropy cutoff, rejecting sites that fail these requirements. A higher first value (60) increases the coverage cutoff (e.g., a site is kept if 60% of all taxa are represented at that site). A higher second value (1.5) increases the entropy cutoff (i.e., entropy values represent the amount of saturation at a site); sites with values higher than 1.5 possess higher saturation and are thus deleted. AliView v1.18 [58] was used to translate to amino acids, check for frame shifts, recognize and remove stop codons, and edit sequencing errors or lone/dubious indels. Because flanking sequences are generally non-coding

and sites have been deemed homologous (see [16]), these flanking sequences, before and after the probe regions, were separated from the exons, then combined and treated together as an independent partition. Due to the filtering steps in the bioinformatics pipeline (i.e., site orthology, and density and saturation evaluation), the flanking partition can be viewed as a SNP supermatrix, where each site is homologous, but uninformative sites, saturated sites, or sites with large amounts missing data have been removed.

Of the 115 bombycoid and two outgroup specimens, 110 were sequenced directly using AHE target capture sequencing, of which 68 were sequenced using the BOM1 and 42 using the LEP1 kit. Seven specimens had their AHE loci probe regions mined from previously sequenced transcriptomes or the *B. mori* genome (Additional file 2: Table S2). These specimens did not have flanking data because of nature of transcriptome data. All specimens were processed using either the 'Bmori' (for LEP1) or 'BOM1' (for BOM1) reference libraries. Previous Breinholt et al., [16] scripts are available in Dryad [46]. Instructions on how to use the pipeline and additional scripts that were not part of Breinholt et al. [16] are provided in Dryad (doi: https://doi.org/10.5061/dryad.5df18fp).

Phylogenetics

Lepidoptera AHE probe sets comprise highly-conserved coding probe regions (i.e., exons) and more variable, generally non-coding flanking regions (e.g., introns or intergenic regions) located on either side of the probe region [16]. We evaluated both nucleotide and amino acid datasets to examine phylogenetic signal and the role that saturation may play in the data (see [19]). With the Breinholt et al. [16] pipeline and scripts [46], three datasets were built for phylogeny inference: 1) AA = an amino acid supermatrix composed of translated probe region loci; 2) Pr + Fl = a probe + flanking supermatrix; and 3) ASTRAL = the individual loci from the Pr + Fl supermatrix, used for individual gene tree inference and then species tree estimation to evaluate the potential effects of deep coalescence. Additional analyses (i.e., probe region-only, as nucleotides) were investigated, but are not reported here because their outcomes did not differ from those reported.

Concatenated supermatrices were assembled using FASconCAT-G v1.02 [59]. Phylogenetic inference was performed in a maximum likelihood (ML) framework using IQ-TREE MPI multicore v1.5.3 [60]. For both nucleotide and amino acid datasets, the '–m TEST' command was used in IQ-TREE to perform a search for the most appropriate model of amino acid or nucleotide substitution. For all inferences, we performed 1000 random addition sequence (RAS) replicates, and 1000 replicates each for both ultrafast bootstraps (UFBS) ('–bb' command) and SH-aLRT tests ('-alrt' command). The SH-like approximate

likelihood ratio test (SH-aLRT) estimates branch support values that have been shown to be as conservative as the commonly used non-parametric bootstrap values [61]. SH-aLRT and bootstrap values tend to agree for data sets with strong phylogenetic signal (i.e., datasets with loci that are sufficiently large in number of bases, and tips that share sufficient divergence between sequences). Disagreements in branch support are thought to arise as a consequence of small sample size, insufficient data, or saturated divergence levels (see [62]). We classified nodes as "robust" if they were recovered with support values of UFBS ≥95 and SH-aLRT ≥80 ([61, 62]).

Because concatenation can be misleading when there are high levels of incomplete lineage sorting or deep coalescence [63], we assessed the impact of potential gene-tree discordance ([64–66]) by inferring a phylogeny for each individual locus, using IQ-TREE under the same parameters as above (i.e., probe and single partition of flanking loci were modeled by site). Species tree estimation was performed in ASTRAL-III [67]. ASTRAL is a computationally efficient and statistically consistent (under the multi-species coalescent) nonparametric method that takes input gene trees and estimates a highly accurate species tree, even when there is a high level of incomplete lineage sorting (or deep coalescence) [68]. The use of ASTRAL is also an informative "data exploration" exercise with phylogenomic datasets, providing valuable information regarding the level of general tree discordance across your set of gene trees, and the potential presence of incomplete lineage sorting/deep coalescence that should be investigated further. To evaluate node support on the species tree, we used the ASTRAL support values (ASV) – local posterior probabilities that are more precise than evaluating bootstrap values across a set of input trees [69]. ASTRAL support values were determined to be "robust" if nodes were recovered with local posterior probabilities ≥0.95. All pipeline steps and phylogenomic analyses were conducted on the University of Florida HiPerGator HPC (http://www.hpc.ufl.edu/). All alignment FASTA files, loci information, partition files, tree files, and other essential data files used for phylogenetic inference are available on Dryad (doi:https://doi.org/10.5061/dryad.5df18fp).

Rogue taxon & outlier locus analyses

We investigated whether our molecular data included rogue taxa or outlier loci that were potentially influencing our phylogenetic results. A rogue taxon analysis was carried out using the online version of RogueNaRok ([70, 71]), http://rnr.h-its.org/, on the 1000 ultrafast bootstrap trees and the consensus tree from the Pr + Fl supermatrix. Outlier taxa and loci analyses were carried out using Phylo-MCOA v.1.4 (PMCoA) [72] in R (on the 650 gene trees used in the ASTRAL analysis). No rogue taxa or loci were found, and therefore we did not prune any taxa or loci from subsequent analyses.

Diversification rate analyses

As an initial investigation into why some bombycoid lineages are more diverse than others, we examined and quantified how diversification rates (the interplay between speciation and extinction) have changed over time. Simply calculating species diversity per clade and assuming extant diversity is a true indicator of increases in diversification rate could produce significant biases in one's interpretations due to some charismatic lineages receiving more taxonomic effort than their "boring" sister lineages (see [73–76]). We therefore applied BAMM [77] and 'BAMM-tools' [78] to infer the number and location of macroevolutionary rate shifts across our phylogeny, and visualize the 95% credible set of shift configurations.

Relative to other arthropod groups, the fossil record of Lepidoptera, especially of the Bombycoidea, is poor [79]. Prior lepidopteran studies that have included fossils in dating analyses have been scrutinized for incorrect fossil identification or placement on the phylogeny [80]. Because the available Bombycoidea fossils are limited and the important characters needed for accurate placement on the phylogeny are difficult to discern, we decided not to conduct a dating analysis for this study. Additionally, we decided against using secondary calibrations from previous studies, as these can be plagued with methodological inaccuracies (e.g., incorrect temporal placement of fossils, inappropriate use of priors, etc.). Instead, the ML best tree was converted into a relative-rate scaled ultrametric tree using the 'chronopl' command in the R package 'ape' [52]. This approach produces a tree whose branches are scaled to evolutionary rates, not a dated tree, and provides a way to understand evolutionary changes over relative "time" of the group being investigated.

For the first time, quantifiable rates of diversification were calculated for the Bombycoidea. This is important, because whether a taxonomic group possesses more described species than another related lineage, does not mean they have diversified "more". A larger number of described species could simply be due to the taxonomic effort, a well-known bias in bombycoids, with the Sphingidae and Saturniidae representing the charismatic groups on which most bombycoid taxonomists have historically worked. To account for non-random missing speciation events, we quantified the percentage of taxa sampled within each family and incorporated these in the form of branch-specific sampling fractions. Sampling fractions were based on the updated superfamily numbers calculated from Kitching et al. [1]. Informed priors, based on our sampling and phylogeny, were determined using

'setBAMMpriors' in BAMMtools. The MCMC chain was run for 100 million generations, sampling every 1000 generations. Convergence diagnostics was assessed using the R package 'coda' [81]. The first 20% of runs were discarded as burn-in.

Results

Our abdomen soaking approach proved relatively consistent with high enough yield of target DNA to proceed with AHE sequencing. Datasets constructed for phylogeny inference contained the following number of loci, sequence length, and model specifications: 1) AA = 579 loci, 48,456 amino acid residues, modeled by locus; 2) Pr + Fl = 649 probe loci + one flanking locus (261,780 bp), each probe locus was modeled by site, flanking data was maintained as a single partition and modeled by site; and 3) ASTRAL = 650 loci, each locus modeled by site. Differences in loci number between the AA dataset and the Pr + Fl dataset are the result of loci being removed from the AA phylogenetic inference due to a lack of variation at the amino acid level for those loci (i.e., the probe regions were highly conserved with no variation across taxa).

Increasing the number of loci and taxon sampling significantly improved our understanding of Bombycoidea relationships. The inferred relationships are generally consistent across the three phylogenetic inferences that we performed (AA, Pr + Fl, ASTRAL), with all major backbone relationships robustly supported, and all bombycoid families sensu Zwick [20] and Zwick et al. [13] were recovered as monophyletic. Due to the methodological approach (i.e., the treatment of different data types) and the more biologically realistic and parsimonious explanation of the topology (see Systematics section in Additional file 10), our preferred phylogeny is the tree generated from the "probe + flanking" (Pr + Fl) dataset (Fig. 1). All family-level placements of genera sensu Kitching et al. [1] were supported, with the exception of *Arotros* Schaus, a genus long considered to be an epiine bombycid [82]. *Arotros* is clearly nested within Apatelodidae in our tree (Fig. 1; Additional file 3: Figure S1) and is hereby transferred to Apatelodidae (see Additional file 10 for additional details on this taxonomic change). The family-level rearrangements of Zwick [20] and Zwick et al. [13] were also recovered in our phylogenetic results: the placement of taxa formerly classified in the "Lemoniidae" (e.g., Lemaire & Minet [23]) are recovered within Brahmaeidae; the Apatelodidae are distinct from the Bombycidae and Phiditiidae; and the broader concept of Endromidae, which includes taxa formerly placed within the "Mirinidae" and Bombycidae (e.g., Lemaire & Minet [23]), are recovered as monophyletic.

We find broad congruence with the major groupings designated by Zwick [20], though internal relationships within these groups did not exactly match previously published trees. Historically, the most problematic familial placement in the superfamily has been the Bombycidae sensu stricto. Phylogenetic studies that were based on a handful of gene regions (e.g., [13, 18, 20]), placed this family either as sister to the Saturniidae or to the Sphingidae (reviewed in [19]), albeit without strong support for either. Our study clearly places the Bombycidae as the sister lineage to the Saturniidae + Sphingidae (the 'SBS' group – coined by Zwick et al. [13]), as seen in trees from the AA and Pr + Fl datasets, but not in the ASTRAL tree (Additional files 7, 8 and 9) – an outcome that mirrors traditional Sanger sequencing studies based on few loci, where individual gene trees can lack the phylogenetic signal of supermatrices. The 'CAPOPEM' group (Carthaeidae, Anthelidae, Phiditiidae, and Endromidae – coined by Regier et al. [11]) is recovered in all three analyses, although in the ASTRAL inference this group is nested within the clade containing the Sphingidae, Bombycidae and Saturniidae (Additional file 4: Figure S2). Interfamilial relationships within the CAPOPEM clade were not robustly supported in Regier et al. [11] or Zwick et al. [13], but our AHE-based trees confidently solidify relationships within this group. Zwick et al. [13] recovered the Old World Endromidae as sister to the Australian Anthelidae + (Australian Carthaeidae + Neotropical Phiditiidae), however, our Pr + Fl analysis supports the Anthelidae as sister to Endromidae + (Carthaeidae + Phiditiidae) instead. Effectively, Anthelidae and Endromidae swap places within CAPOPEM when comparing Zwick et al. [13] and our Pr + Fl phylogeny. Conversely, our AA results illustrate yet a different picture, in which Phiditiidae are sister to Anthelidae + (Carthaeidae + Endromidae), mirroring Regier et al. [11]. This discordance is important to note considering the disjunct geographic distribution of these families. Among these families, only Phiditiidae is found in (and endemic to) the New World, with each of the other families being restricted to the Old World. Lastly, while intrafamilial relationships in the SBS and CAPOPEM groups change depending on which of our datasets are used to infer the phylogeny (Fig. 1, Additional file 3: Figure S1, Additional file 4: Figure S2, Additional file 5: Figure S3 and Additional files 7, 8, and 9), the 'BALE' group (Brahmaeidae, Apatelodidae, and Eupterotidae – coined by Regier et al. [11]) is recovered in all three analyses.

When we investigate intrafamilial relationships, the Bombycidae (excluding *Arotros*, mentioned above; see Additional file 10) possesses distinct and monophyletic subfamilies Bombycinae (Old World) and Epiinae (New World; Fig. 1; Additional file 3: Figure S1). Within the Eupterotidae, the Striphnopteryginae and Janinae are monophyletic, but others are not (Fig. 1; Additional file 3: Figure S1). Within the Saturniidae, the Pr + Fl and AA topologies are similar to Regier et al. [32] (Fig. 1 &
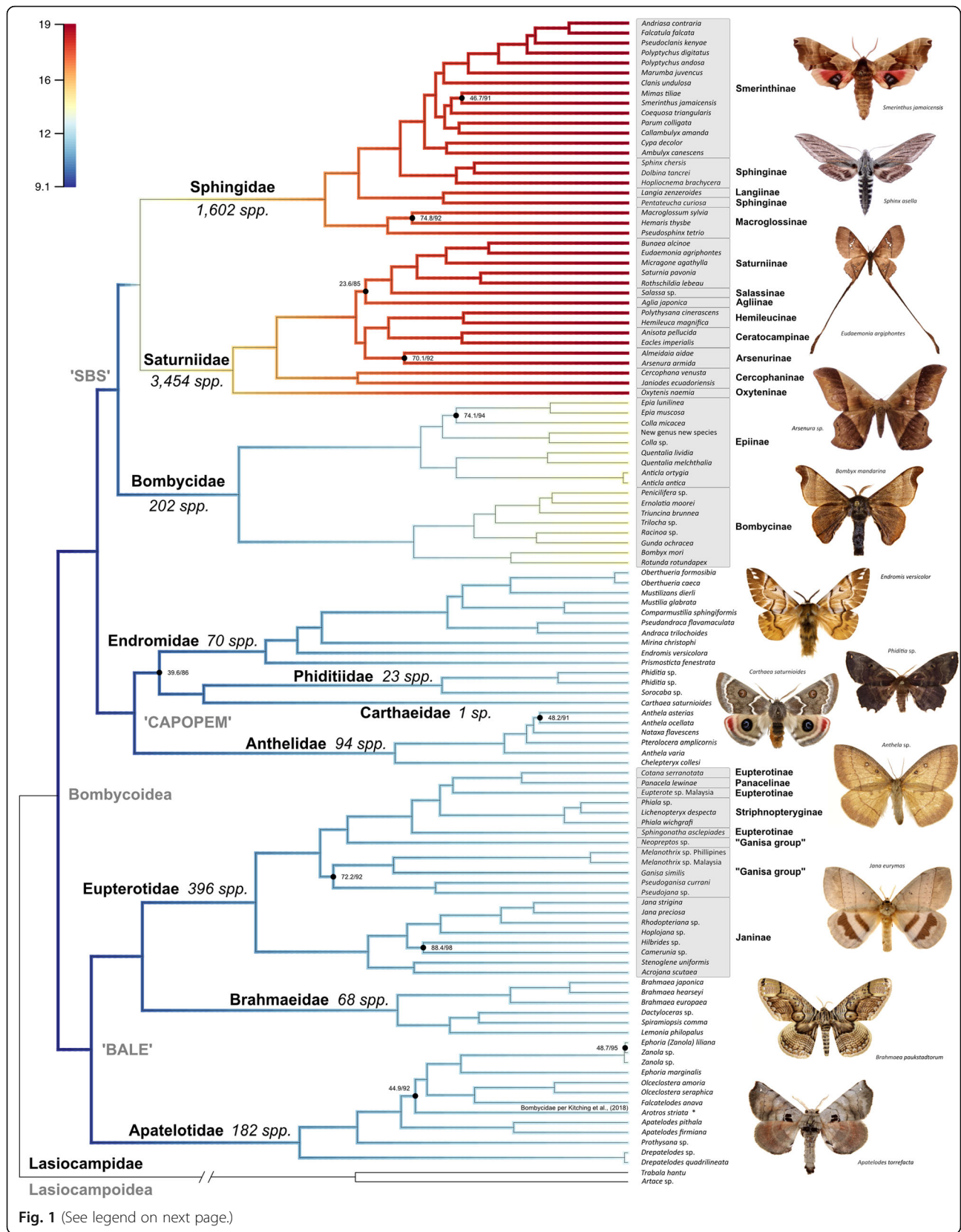
Fig. 1 (See legend on next page.)

Additional file 5: Figure S3), where the Oxyteninae are sister to the rest of the family, followed by the Cercophaninae. Arsenurinae is the sister lineage to the Hemileucinae and the Ceratocampinae, while Salassinae is the sister lineage to the Saturniinae. A major difference in subfamily relationships is the placement of Agliinae, which was either sister to the Salassinae + Saturniinae (Pr + Fl) or to the Arsenurinae (AA, also [13, 32]). The ASTRAL inference is quite different, placing the Ceratocampinae as the sister lineage to a clade containing the Agliinae, Arsenurinae + Hemileucinae, and a Saturniinae + Salassinae clade (Additional file 5: Figure S3), albeit with low branch support. Within the Sphingidae, our topologies (both Pr + Fl and AA) are largely congruent, but differ slightly from Kawahara et al. [83], Zwick et al. [13], and Kawahara and Barber [34]. Smerinthinae, which now excludes *Langia* (Kitching et al. [1]) is monophyletic and the Sphinginae are polyphyletic (Fig. 1; Additional file 3: Figure S1). This outcome is robustly supported in the Pr + Fl and AS-TRAL trees, although internal relationships slightly differ between the two trees. The placement of the Macroglossinae, Langiinae, and Sphinginae are robustly supported in the AA and Pr + FI tree, but low internal node support obfuscates relationships within Smerinthinae.

Among the other families, a number of subfamilies and genera, as defined by the most current classification of the Bombycoidea [1], are not monophyletic (Fig. 1; Additional file 3: Figure S1). Within Eupterotidae, the Eupterotinae and the "Ganisa group" are not monophyletic. The Eupterotinae are rendered paraphyletic due to the placements of Panacelinae and Striphnopteryginae. The Striphnopteryginae genus *Phiala* is paraphyletic with respect to *Lichenopteryx*, and the Ganisa group is polyphyletic due to the traditional inclusion of the genus *Neopreptos* [84]. Within the Anthelidae, *Anthela* is paraphyletic due to the placement of *Nataxa* and *Pterolocera*, a finding congruent with Zwick [20]. Within Sphingidae, Sphinginae is polyphyletic due to the placement of *Pentateucha* as sister to the Langiinae, and the Smerinthinae genus *Polyptychus* is paraphyletic.

A number of factors can lead to the appearance of a taxonomic group being more "diverse" than other sister lineages. For example, when simply looking at numbers of described species, taxonomic bias in interest and effort could substantially affect our understanding. This is why it is essential to formally test diversification rates. While there are more described species of Saturniidae and Sphingidae than the rest of the Bombycoidea, that doesn't mean those two families are necessarily more evolutionarily diverse. For the first time, actual quantifiable rates of diversification were calculated for the Bombycoidea, by evaluating how the interplay between speciation and extinction has changed over relative time. At the family level, a major diversification rate shift occurs along the lineage leading to the Saturniidae and Sphingidae (Fig. 1). One important aspect of our results relates to the way we applied the sampling fraction (i.e., at the family level). By doing this, our results should be interpreted as a generalized look at rate shifts at the family level, lacking information for the more derived branches. The effective sample size of the log-likelihood was 8475.169 and the effective sample size of the number of shift events present in each sample was 18,141.4. The 95% credible set of rate shift configurations sampled with BAMM can be seen in Additional file 6: Figure S4. Unfortunately, due to the poor Bombycoidea fossil record and the limited computational approaches currently available to adequately estimate extinction rates, we are unable to discern whether the differences that exist are due to increases or decreases in speciation, although species diversity numbers, based on the classification of Kitching et al. [1] agree with our diversification rate shift results. The number of described species for each of the ten bombycoid families, according to Kitching et al. [1] is: Anthelidae (94 spp.); Apatelotidae (182 spp.); Bombycidae (202 spp.); Brahmaeidae (68 spp.); Carthaeidae (1 sp.); Endromidae (70 spp.); Eupterotidae (396 spp.); Phiditiidae (23 spp.); Saturniidae (3454 spp.); Sphingidae (1602 spp.).

Discussion

Within the past decade, a few studies using molecular sequence data have attempted to resolve the phylogenetic relationships of the Bombycoidea ([11, 13, 20]). None corroborated the earlier morphology-based hypotheses ([15, 23]). In order to establish a backbone for future evolutionary, comparative, and taxonomic studies, we sampled exemplars from all major lineages in the superfamily and used Anchored Hybrid Enrichment (AHE) phylogenomics to provide a robust phylogeny of the superfamily based on the largest taxonomic and molecular sampling to date.

To achieve our representative taxonomic coverage, we included many samples from pinned specimens in natural history collections. We modified previous DNA extraction protocols to increase DNA yield for high-throughput sequencing and modified the AHE probe set developed by

Breinholt et al. [16] to more efficiently recover phylogenetically-informative loci within the Bombycoidea. To allow more flexibility in the use of the data (e.g., integration of these samples with "legacy" datasets and CO1 for species identification), we added a selection of traditional Sanger-sequenced loci. To assist future usage of the BOM1 and other AHE probe sets, we expanded the bioinformatic protocol of Breinholt et al. [16] to include step-by-step instructions for running the bioinformatics pipeline.

Our phylogeny largely reinforces the results of earlier molecular works, phylogenomics brings into focus some of the long problematic relationships, while also identifying important topological regions where some of the traditionally-defined subfamilies and genera are not recovered as monophyletic. When comparing with trees published in previous works, topological discordances are likely the product of increased locus sampling, which provided significantly more phylogenetic information, as well as morphological homoplasy or convergence that likely obscured the true placement of certain taxonomic groups. For example, our results provide a well-supported placement of the historically troublesome family Bombycidae sensu stricto, as the sister lineage to Saturniidae + Sphingidae. An earlier phylogenomic study [19] had provided some evidence in support of this relationship, but its taxon sampling was very limited.

Our study highlights how morphological convergence in Bombycoidea has confused our understanding of their evolution. For example, the bombycid genus *Rotunda*, endemic to the Old World, and the apatelodid genus *Arotros*, endemic to the New World, are clearly unrelated (Fig. 1; Additional file 3: Figure S1). However, they are both astonishingly similar in body size, wing shape, and phenotypic appearance (see Additional file 10: Systematics). Such results imply that there may be adaptive advantages to evolving particular wing shape and size, as found in other bombycoid lineages such as in the Saturniidae [30]. Furthermore, within the CAPOPEM clade, we see a biogeographical distribution that, based on the AA and ASTRAL topologies, appears to be a more biologically realistic and parsimonious explanation (than the Pr + Fl topology) of the evolutionary history due to the placement of the Neotropical Phiditiidae as sister to the remaining and Old World CAPOPEM families (Additional file 4: Figure S2). That said, the potential for biases in phylogenomic inference is an area of important investigation and one where we are beginning to better understand the forces behind these biases (see [85, 86]). For example, incongruent topologies can be obtained, with high "support", simply due to side effects of the phylogenetic inference algorithms and the historically-used methods for evaluating nodal support (i.e., bootstrapping - where simply the addition of more data ends up providing high, but misleading, support that is not truly indicative of the

phylogenetic information within the data). Because the development and implementation of improved evaluations of support are in their relative infancy, we suggest that a level of caution be taken with phylogenomic topology inference herein. In the future, increased taxon sampling and investigation of novel support methods will likely help bring better resolution to both the backbone and internal bombycoid relationships. These types of findings, wherein different dataset types answer different questions, highlight the importance of evaluating phylogenetic data in different ways because phylogenetic signal could be hiding in phylogenomic datasets.

From an evolutionary viewpoint, one of the most interesting results came from the first attempt to quantify the diversification rates across the Bombycoidea, in particular, the dramatic shift in diversification rates leading to the Sphingidae and Saturniidae lineages. While maybe not "surprising", because of the number of species described, this had never been quantified before in this group. We know that simple numbers of described species does not mean necessarily that they have diversified "more" than other closely related lineages. A number of factors could cause one taxonomic group to appear as if there are more species than another related lineage. Perhaps the most important of these is simply due to the taxonomic effort that has historically been applied to a group, a bias that can certainly be found in the bombycoids, with Sphingidae and Saturniidae representing highly charismatic groups where most bombycoid taxonomists have worked.

These findings are fascinating because the Sphingidae and Saturniidae have contrasting life-history strategies, such as larvae feeding on different kinds of food plants, adults having the ability to nectar source [36], and anti-bat strategies ([30, 33–35]) which are reflected in their morphology (e.g., body sizes and wing shapes) and behavior, including flight speed and maneuverability [36]. In an "arms race", as has been shown between moths and bats, bat predation selectively removes unfit lineages from the environment, thus increasing the speed of evolution of these surviving lineages. As new traits evolve that can be used to effectively evade the predator, subsequent release from the predatory pressure provides the opportunities to diversify ecologically and behaviorally. Because the Bombycoidea lack sufficiently informative fossils that can (or should) be used in time-calibrating a phylogeny, we did not conduct a divergence time estimation analysis in the present study, although from previous estimations, we have a general evolutionary timeframe of when the origin of the SBS group occurred. This origin has been postulated to have occurred either approximately 50 mya [87], with Sphingidae originating soon thereafter [34], or approximately 70–80 mya [88] – a difference that could potentially impact the results of our diversification analysis. Insectivorous bats

are thought to have originated roughly 60 mya, and the diversification of Sphingidae and Saturniidae around that time suggests that the incredible taxonomic diversity within these two families could be in part due to bat-related selection pressures resulting in diverse anti-bat traits. Importantly, these diversification rate findings should be considered preliminary, and that with denser sampling across the Bombycoidea, researchers will likely uncover more distinct and informative shifts.

The hawkmoth and silk moth evolutionary story is certainly more complex than simply reflecting their interactions with bats. Although we postulate that the differences in diversification rates are correlated with bat predation, it is possible that these rate shifts are due to other factors, such as ecological specialization or shifts in host plant usage, both as larvae and adults. Amassing and collating behavioral and ecological datasets for the tips of the Bombycoidea Tree of Life for macroevolutionary comparative investigations is essential to furthering our understanding of this diverse, global superfamily, as well as understanding how bats, ecological traits, and/or biogeographical history may or may not have shaped their diversity. At this time, large trait datasets do not exist for these groups, but are currently being worked on to be included in future trait-dependent diversification analyses with much more complete sampling, at the genus and species level across these families, to truly explore the drivers of Bombycoidea diversity. This research establishes some hypotheses to be further tested when more complete sampling of the Bombycoidea has been completed, and robust trait datasets have been collected.

Conclusions

Our study finally brings into focus long problematic bombycoid relationships and establishes a backbone for future evolutionary, comparative, and taxonomic studies. Our modified DNA extraction protocol allows Lepidoptera specimens to be readily sequenced from pinned natural history collections and highlights the flexibility of AHE to generate genomic data from a wide range of museum specimens, both age and preservation method. By allowing researchers to tap into the wealth of biological data residing in natural history collections around the globe, these types of methodologies (e.g., DNA from museum specimens and targeted sequencing capture) will provide the opportunities for us to continually add to our understanding of Lepidoptera and Bombycoidea evolution, as well as refine our understanding of relationships across the Tree of Life.

Additional files

Additional file 1: Table S1. List of the loci names in the modified Bombycoidea-specific AHE probe set "BOM1". The first column lists the name of the locus; the second column explains what the locus was named in the "LEP1" kit or the name of the "legacy" or vision-related loci; the third column lists the length (bp) of each locus. (XLSX 25 kb)

Additional file 2: Table S2. List of the total taxa included in the phylogenetic analyses. Specimen names refer to tip names and include the specimen accession number and preliminary taxonomy. These specimen names can be used to identify a specimen's specific molecular data in the supplemental data. Other pieces of information included: taxonomy (family, subfamily, tribe, genus, and species – per Kitching et al., (12)); storage method of the tissues; whether the data came from an AHE probe set, transcriptome, or genome; and collecting date. (XLSX 16 kb)

Additional file 3: Figure S1. Maximum likelihood tree of Bombycoidea, based on 650 AHE loci. All nodes are supported by ≥95% bootstrap values unless otherwise noted. Major taxonomic groups such as families, subfamilies, and tribes are labeled. Red circles at the tips correspond to genera that are not monophyletic. Black boxes around the tips correspond to non-monophyletic subfamilies. (TIFF 1188 kb)

Additional file 4: Figure S2. Interfamilial relationships of the 'CAPOPEM' group. These relationships change depending on the data used (Pr + Fl or AA) to infer the phylogeny and the phylogenetic inference (supermatrix or ASTRAL). Values at the nodes of the Pr + Fl or AA trees indicate SH-aLRT/UFBS support, or ASV for the ASTRAL analysis. (TIFF 1294 kb)

Additional file 5: Figure S3. Interfamilial relationships of the Saturniidae. These relationships change depending on the data used (Pr + Fl or AA) to infer the phylogeny and the phylogenetic inference method (supermatrix or ASTRAL). Values at the nodes of the Pr + Fl or AA trees indicate SH-aLRT/UFBS support, or ASV for the ASTRAL analysis. (TIFF 1224 kb)

Additional file 6: Figure S4. The four most common 95% credible set of rate shift configurations sampled with BAMM. Branch color indicates the estimated diversification rate, with warmer colors representing lineages with higher rates. Major taxonomic groups with shifts are labeled. Photographs correspond to the major lineages (i.e., the SBS group) with diversification shifts in the phylogeny. (TIFF 1055 kb)

Additional file 7: phylogeny 1 The Pr + Fl phylogeny. (TXT 8 kb)

Additional file 8: phylogeny 2. The AA phylogeny. (TXT 8 kb)

Additional file 9: phylogeny 3 The ASTRAL phylogeny. (TXT 7 kb)

Additional file 10: Supplemental Information. The Supplemental Information contains information regarding where the specimens that were used for this study are housed, as well as the DNA extraction protocol of dried museum specimens, and the Systematics section. (PDF 179 kb)

Abbreviations

AA: Amino acid supermatrix; AHE: Anchored Hybrid Enrichment; ASV: ASTRAL support values; BALE: Brahmaeidae, Apatelodidae, and Eupterotidae; BOM1: The Bombycoid-specific AHE probe set; CAPOPEM: Carthaeidae, Anthelidae, Phiditiidae, and Endromidae; LEP1: The Lepidoptera-wide AHE probe set; MGCL: McGuire Center of Lepidoptera and Biodiversity; ML: Maximum likelihood; Pr + Fl: Probe + flanking supermatrix; RAS: Random addition sequence replicates; SBS: Sphingidae, Bombycidae and Saturniidae; SH-aLRT: SH-like approximate likelihood ratio test; UCE: Ultraconserved Elements; UFBS: Ultrafast bootstraps

Authors' contributions

Funding

Availability of data and materials

The datasets supporting the results of this article are available in Dryad Data Repository (doi:https://doi.org/10.5061/dryad.5df18fp). The supplemental information includes four figures, one table, plus the full museum specimen DNA extraction protocol, systematic information on Bombycoidea families, bioinformatics pipeline scripts and instructions, alignment FASTA files (nucleotides, amino acids, supermatrices, and individual loci), partition files, tree files, and other essential data files that were inputted or outputted from the phylogenetic inference and analysis. Previous Breinholt et al., [16] bioinformatics scripts are available in Dryad [46].

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA. [2]Department of Entomology, Plant Pathology & Nematology, University of Idaho, Moscow, ID 83844, USA. [3]Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK. [4]RAPiD Genomics, 747 SW 2nd Avenue #314, Gainesville, FL 32601, USA. [5]Australian National Insect Collection, CSIRO, Clunies Ross St, Acton, ACT, Canberra 2601, Australia. [6]Department of Natural Sciences, Middlesex University, The Burroughs, London NW4 4BT, UK. [7]Department of Biological Sciences, Boise State University, Boise, ID 83725, USA.

References

1. Kitching I, Rougerie R, Zwick A, St Laurent R, Ballesteros Mejia L, Kawahara A, et al. A global checklist of the Bombycoidea (Insecta: Lepidoptera). BDJ. 2018;6:e22236–13.
2. Roe AD, Weller SJ, Baixeras J, Brown J, Cummings MP, Davis DR, et al. Evolutionary framework for Lepidoptera model systems. In: Goldsmith M, Marec F, editors. Molecular biology and genetics of the Lepidoptera; 2010. p. 1–24.
3. Gregory D. Hawkmoth pollination in the genus *Oenothera*. Aliso J Syst Evol Bot. 1963;5:357–84.
4. Haber WA, Frankie GW. A tropical hawkmoth community- Costa Rican dry Forest Sphingidae. Biotropica. 1989;21:155–72.
5. Nilsson LA, Jonsson L, Rason L, Randrianjohany E. Monophily and pollination mechanisms in *Angraecum arachnites* Schltr. (Orchidaceae) in a guild of long-tongued hawk-moths (Sphingidae) in Madagascar. Biol J Linn Soc. 1985;26:1–19.
6. Nilsson LA, Rabakonandrianina E, Pettersson B. Exact tracking of pollen transfer and mating in plants. Nature. 1992;360:666–8.
7. Wasserthal LT. The pollinators of the Malagasy star orchids *Angraecum sesquipedale, A. sororium* and *A. compactum* and the evolution of extremely long spurs by pollinator shift. Botanica Acta. 1997;110:343–59.
8. Nilsson LA. Deep flowers for long tongues. Trends Ecol Evol. 1998;13:259–60.
9. Raguso RA, Willis MA. Synergy between visual and olfactory cues in nectar feeding by naive hawkmoths, *Manduca sexta*. Anim Behav. 2002;64:685–95.
10. Beck J, Kitching IJ, Linsenmair KE. Effects of habitat disturbance can be subtle yet significant: biodiversity of hawkmoth-assemblages (Lepidoptera: Sphingidae) in Southeast-Asia. In: Hawksworth DL, Bull AT. (eds) Arthropod Diversity and Conservation. Topics in Biodiversity and conservation, vol. 1. Dordrecht: Springer; 2006.
11. Regier JC, Cook CP, Mitter C, Hussey A. A phylogenetic study of the "bombycoid complex" (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny. Syst Entomol. 2008;33:175–89.
12. Meusemann K, Reumont v BM, Simon S, Roeding F, Strauss S, Kück P, et al. A phylogenomic approach to resolve the arthropod tree of life. Mol Biol Evol. 2010;27:2451–64.
13. Zwick A, Regier JC, Mitter C, Cummings MP. Increased gene sampling yields robust support for higher-level clades within Bombycoidea (Lepidoptera). Syst Entomol. 2011;36:31–43.
14. Kim MJ, Kang AR, Jeong HC, Kim K-G, Kim I. Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae). Mol Phylogenet Evol. Elsevier Inc. 2011;61:436–45.
15. Minet J. The Bombycoidea: Phylogeny and higher classification (Lepidoptera, Glossata). Entomologica Scandinavica. 1994;25:63–88.
16. Breinholt JW, Earl C, Lemmon AR, Lemmon EM, Xiao L, Kawahara AY. Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for anchored phylogenomics. Syst Biol. 2018;67:78–93.
17. Timmermans MJTN, Lees DC, Simonsen TJ. Towards a mitogenomic phylogeny of Lepidoptera. Mol Phylogenet Evol. Elsevier Inc. 2014;79:169–78.
18. Wang X, Chen Z-M, Gu X-S, Wang M, Huang G-H, Zwick A. Phylogenetic relationships among Bombycidae s.l. (Lepidoptera) based on analyses of complete mitochondrial genomes. Syst Entomol. John Wiley & Sons, Ltd. 2018;20:537–9.
19. Breinholt JW, Kawahara AY. Phylotranscriptomics: saturated third codon positions radically influence the estimation of trees based on next-gen data. Genome Biol Evol. 2013;5:2082–92.
20. Zwick A. Molecular phylogeny of Anthelidae and other bombycoid taxa (Lepidoptera: Bombycoidea). Syst Entomol. 2008;33:190–209.
21. Minet J. Tentative reconstruction of the ditrysian phylogeny (Lepidoptera, Glossata). Entomologica Scandinavica. 1991;22:69–95.
22. Mutanen M, Wahlberg N, Kaila L. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proc Biol Sci. 2010; 277:2839–48.
23. Lemaire C, Minet J. 18. The Bombycoidea and their relatives. In: Kristensen NP, editor. Lepidoptera, Moths and Butterflies. vol. 1. Evolution, Systematics, and Biogeography. Handbook of Zoology. Berlin: Walter de Gruyter; 1998. p. 322–53.
24. Regier JC, Mitter C, Zwick A, Bazinet AL, Cummings MP, Kawahara AY, et al. A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). Moreau CS, editor. PLoS ONE. 2013;8: e58568–23.
25. Regier JC, Zwick A, Cummings MP, Kawahara AY, Cho S, Weller S, et al. Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. BMC evolutionary biology. BioMed Central. 2009;9:280–21.
26. Espeland M, Breinholt J, Willmott KR, Warren AD, Vila R, Toussaint EFA, et al. A comprehensive and dated phylogenomic analysis of butterflies. CURBIO. Elsevier Ltd. 2018;28:770–5.
27. Johns CA, Toussaint EFA, Breinholt JW, Kawahara AY. Origin and macroevolution of micro-moths on sunken Hawaiian islands. Proc Biol Sci. 2018;285:20181047–10.
28. St Laurent RA, Hamilton CA, Kawahara AY. Museum specimens provide phylogenomic data to resolve relationships of sack-bearer moths (Lepidoptera, Mimallonoidea, Mimallonidae). Syst Entomol. Wiley/Blackwell (10.1111). 2018;43:729–61.
29. Toussaint EFA, Breinholt JW, Earl C, Warren AD, Brower AVZ, Yago M, et al. Anchored phylogenomics illuminates the skipper butterfly tree of life. BMC Evol Biol. 2018;18:1–11.

30. Rubin JJ, Hamilton CA, McClure CJW, Chadwell BA, Kawahara AY, Barber JR. 1028 The evolution of anti-bat sensory illusions in moths. Science advances. 1029 American association for the. Adv Sci. 2018;4:EAAR7428. p. 1-9.

31. Lemmon AR, Emme SA, Lemmon EM. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst Biol. 2012;61:727–44.

32. Regier JC, Grant MC, Mitter C, Cook CP, Peigler RS, Rougerie R. Phylogenetic relationships of wild silkmoths (Lepidoptera: Saturniidae) inferred from four protein-coding nuclear genes. Syst Entomol. 2008;33:219–28.

33. Barber JR, Kawahara AY. Hawkmoths produce anti-bat ultrasound. Biol Lett. 2013;9:20130161.

34. Kawahara AY, Barber JR. Tempo and mode of antibat ultrasound production and sonar jamming in the diverse hawkmoth radiation. Proc Natl Acad Sci. 2015;112:6407–12.

35. Barber JR, Leavell BC, Keener AL, Breinholt JW, Chadwell BA, McClure CJW, et al. Moth tails divert bat attack: evolution of acoustic deflection. Proceedings of the National Academy of Sciences. Natl Acad Sci. 2015; 112:2812–6.

36. Janzen D. Two ways to be a tropical big moth: Santa Rosa saturniids and sphingids. Oxf Surv Evol Biol. 1984;1:140.

37. Tammaru T, Haukioja E. Capital breeders and income breeders among Lepidoptera: consequences to population dynamics. Oikos. 1996;77:561–4.

38. Kawahara AY, Breinholt JW. Phylogenomics provides strong evidence for relationships of butterflies and moths. P Roy Soc B-Biol Sci. 2014;281: 20140970.

39. Jones KE, Bininda-Emonds ORP, Gittleman JL. Bats, clocks, and rocks: diversification patterns in Chiroptera. Evolution. 2005;59:2243–55.

40. Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ. A molecular phylogeny for bats illuminates biogeography and the fossil record. Science. 2005;307:580–4.

41. Miller-Butterworth CM, Murphy WJ, O'Brien SJ, Jacobs DS, Springer MS, Teeling EC. A family matter: conclusive resolution of the taxonomic position of the long-fingered bats, *Miniopterus*. Mol Biol Evol. 2007;24: 1553–61.

42. Shi JJ, Rabosky DL. Speciation dynamics during the global radiation of extant bats. Evolution. 2015;69:1528–45.

43. Lei M, Dong D. Phylogenomic analyses of bat subordinal relationships based on transcriptome data. Sci Rep. 2016;6:27726.

44. Barber JR, Conner WE. Acoustic mimicry in a predator-prey interaction. P Natl Acad Sci Usa. 2007;104:9331–4.

45. Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, et al. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). Science. 2004;306: 1937–40.

46. Breinholt JW, Earl C, Lemmon AR, Lemmon EM, Xiao L, from KAYD. Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for anchored Phylogenomics. Dryad Digital Repository. 2017. https:// doi.org/10.5061/dryad.rf7g5.2.

47. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(22):1312–13. https:// doi.org/10.1093/bioinformatics/btu033.

48. Mayrose I, Graur D, Ben-Tal N, Pupko T. Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. Mol Biol Evol. 2004;21:1781–91.

49. Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. New York: Springer; 2005. p. 125–81.

50. López-Giráldez F, Townsend JP. PhyDesign: an online application for profiling phylogenetic informativeness. BMC Evol Biol. BioMed Central Ltd. 2011;11:152.

51. Heibl C. PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software packages; 2013.

52. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289–90.

53. Kawahara AY, Plotkin D, Hamilton CA, et al. Diel behavior in moths and butterflies: a synthesis of data illuminates the evolution of temporal activity. Org Divers Evol. 2018;18:13. https://doi.org/10.1007/s13127-017-0350-6.

54. Cho S, Epstein SW, Mitter K, Hamilton CA, Plotkin D, Mitter C, et al. Preserving and vouchering butterflies and moths for large-scale museum-based molecular research. PeerJ. 2016;4:e2160.

55. Blaimer BB, Lloyd MW, Guillory WX, Brady SG. Sequence capture and phylogenetic utility of genomic Ultraconserved elements obtained from pinned insect specimens. PLoS ONE, O'Grady P, editor. Public Libr Sci. 2016; 11:e0161531–20.

56. Mikheyev AS, Zwick A, Magrath MJL, Grau ML, Qiu L, Su YN, et al. Museum genomics confirms that the Lord Howe Island stick insect survived extinction. CURBIO. Elsevier Ltd. 2017;27:3157–3161.e4.

57. Ruane S, Austin CC. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. Molecular Ecology Resources. Wiley/Blackwell (10.1111). 2017;17:1003–8.

58. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8. https://doi.org/10.1093/bioinformatics/btu531.

59. Kück P, Longo GC. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front Zool. 2014; 11:81, 1–8.

60. Nguyen L-T, Schmidt HA, Haeseler v A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.

61. Minh BQ, Nguyen MAT, Haeseler v A. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol. 2013;30:1188–95.

62. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21.

63. Mendes FK, Hahn MW. Why concatenation fails near the anomaly zone. Syst Biol. 2017;67:158–69.

64. Maddison W. Gene trees in species trees. Syst Biol. 1997;46:523–36.

65. Slowinski J, Page R. How should species phylogenies be inferred from sequence data? Syst Biol. 1999;48:814–25.

66. Edwards SV. Is a new and general theory of molecular systematics emerging? Evolution. Wiley Online Library. 2009;63:1–19.

67. Zhang C, Sayyari E, Mirarab S. ASTRAL-III: increased scalability and impacts of contracting low support branches. RECOMB International Workshop on Comparative Genomics; 2017. p. 53–75.

68. Mirarab S, Warnow T. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics. 2015; 31:i44–52.

69. Sayyari E, Mirarab S. Anchoring quartet-based phylogenetic distances and applications to species tree reconstruction. BMC Genomics. BioMed Central. 2016;17:101–13.

70. Aberer AJ, Stamatakis A. A simple and accurate method for rogue taxon identification. IEEE International Conference on Bioinformatics and Biomedicine BIBM; 2011. p. 118–22.

71. Aberer AJ, Krompass D, Stamatakis A. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. Syst Biol. 2012;62:162–6.

72. de Vienne DM, Ollier S, Aguileta G. Phylo-MCOA: a fast and efficient method to detect outlier genes and species in phylogenomics using multiple co-inertia analysis. Mol Biol Evol. 2012;29:1587–98.

73. Dial KP, Marzluff JM. Nonrandom diversification within taxonomic assemblages. Syst Zool. 1989;38:26–37.

74. Doyle JA, Donoghue MJ. Phylogenies and angiosperm diversification. Paleobiology. 1993;19:141–67.

75. Sanderson MJ, Donoghue MJ. Shifts in diversification rate with the origin of angiosperms. Science. 1994;264:1590–3.

76. Sanderson MJ, Donoghue MJ. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol Evol. 1996;11:15–20.

77. Rabosky DL. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. Kolokotronis S-O. PLoS One. 2014;9:e89543.

78. Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, et al. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. Kembel S, editor. Methods Ecol Evol. 2014;5:701–7.

79. Sohn JC, Labandeira CC, Davis DR, Mitter C. An annotated catalog of fossil and subfossil Lepidoptera (Insecta: Holometabola) of the world. Zootaxa. 2012:1–132.

80. de Jong R. Fossil butterflies, calibration points and the molecular clock (Lepidoptera: Papilionoidea). Zootaxa. 2017;4270:1–63.

81. Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis for MCMC. R News. 2006;6:7–11.

82. Schaus W. Familie Bombycidae (Epiinae). In: A. Seitz, Kernen. editor. Die Gross-Schmetterlinge der Erde; 1929. p. 675–711.

83. Kawahara AY, Mignault AA, Regier JC, Kitching IJ, Phylogeny MC. Biogeography of hawkmoths (Lepidoptera: Sphingidae): evidence from five nuclear genes. Knapp M, editor. PLoS ONE. 2009;4:e5719.

84. Nässig WA, Oberprieler RG. An annotated catalogue of the genera of Eupterotidae (Insecta, Lepidoptera, Bombycoidea). Senckenberg Biol. 2008; 88:53–80.
85. Scornavacca C, Galtier N. Incomplete Lineage Sorting in Mammalian Phylogenomics. Syst Biol. 2017;5:syw082–9.
86. Dornburg A, Su Z, Townsend JP. Optimal rates for phylogenetic inference and experimental Design in the era of genome-scale data sets. Mueller R, editor. Syst Biol. 2019;68:145–56.
87. Misof B, Liu S, Meusemann K, Peters RS, Donath A. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014;346:763–7.
88. Wahlberg N, Wheat CW, Peña C. Timing and patterns in the taxonomic diversification of Lepidoptera (butterflies and moths). PLoS One. 2013;8:e80875.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.