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

Hamilton CA[1,2*], St Laurent RA[1], Dexter, K[1], Kitching IJ[3], Breinholt JW[1,4], Zwick A[5], Timmermans MJTN[6], Barber JR[7], Kawahara AY[1*]

Institutional Affiliations:
[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 2601, Canberra, 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
[*]Correspondence: hamiltonlab@uidaho.edu (CAH) or kawahara@flmnh.ufl.edu (AYK)

Abstract

The silkmoths and their relatives (Bombycoidea) are an ecologically and taxonomically diverse superfamily that includes some of the most charismatic species of all the Lepidoptera. Despite displaying some of the most spectacular forms and ecological traits among insects, relatively little attention has been given to understanding their evolution and the drivers of their diversity. We heavily sampled (both in taxa and loci) all major lineages of the Bombycoidea, producing a well-supported phylogeny that identified important evolutionary patterns (e.g., morphology, biogeography, and differences in speciation and extinction). Importantly, analysis of diversification rates highlights the stark increases that exist within the Sphingidae (hawkmoths) and Saturniidae (wild silkmoths). We postulate that these rate shifts are due to differences in the intense selective pressures from insectivorous bats. The study also introduces a new Bombycoidea-specific Anchored Hybrid Enrichment (AHE) probe set, a modified DNA extraction protocol for Lepidoptera specimens from natural history collections, and additional information on the existing AHE bioinformatics pipeline. 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, Bombycoidea, Lepidoptera, natural history collections

Background

The Bombycoidea include some of the most charismatic moth species of all the Lepidoptera. This ecologically diverse superfamily comprises ten families, 520 genera, and 6,092 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). Their large body sizes, ease of rearing in captivity and utilization of larval silk has driven human interest in these moths.

A number of bombycoids have become important contributors to human culture, originally as economically important species for either 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), (4), (5), (6), (7), (8), (9), or as indicators in biodiversity and habitat quality assessments (10). Of the 10 families, three contain species that have been used as model organisms (Bombycidae, Saturniidae, and Sphingidae). Unfortunately, the relationships among these three families have remained largely elusive, with Bombycidae considered the sister lineage to either Saturniidae or Sphingidae (11), (12), (13), (14), (15), (16), (17), (18); but see Breinholt and Kawahara (19). Overall, relatively little attention has been given to the group with regards to understanding their evolution.

Despite their charisma and intrigue, the lack a robust phylogeny based on broad taxon sampling is dramatically affecting our ability to answer fundamental questions about the drivers of their diversity. Monophyly of the Bombycoidea has been supported by six morphological synapomorphies (20), but Zwick (21) determined that only two of these were systematically informative: one poorly understood thoracic character (22), and one relating to forewing vein arrangements (A. Zwick, unpublished). Recent molecular studies of bombycoid systematics (11), (13), (21) have resulted in substantial differences from morphology-based phylogenetic hypotheses (20), (22), (23). To date, nearly all molecular studies of the superfamily have included fewer than 20 protein-coding genes for ≤50 species (11), (13), (16), (24). 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 (25), (26), (27), (28), (29), including within the Bombycoidea (25), (30). However, studies that utilize phylogenomics to resolve inter- and intra-familial relationships within the Bombycoidea have been limited by sampling.

Perhaps the most interesting evolutionary aspect of this superfamily is their striking link with bats, their main nocturnal predators (30), (31), (32), (33), (34). Bombycoids possess a number of traits that appear to be anti-bat adaptations in response to echolocating bats, a lineage thought to have arisen approximately 60 million years ago (35), (36), (37), (38), (39). These nightly battles between bats and moths drove a predator/prey “arms race” that produced anti-bat strategies such as hearing organs (40), (41), (42), (43), ultrasound producing organs capable of jamming bat sonar (33), (44), evasive flight strategies (both in aerial maneuvers and temporal partitioning of activity (45), and a variety of different anti-bat wing morphologies (30), (46).

The most spectacular examples of these anti-bat strategies can be found in the Saturniidae and Sphingidae – the two lineages that harbor the majority of described bombycoid species (30), (32), (33), (47). The divergent life-history strategies of these two families (31), (48) has likely played a major role in driving their diversity by exposing these lineages to differing

predatory selective forces. For example, the majority of hawkmoths feed as adults, seeking out nectar resources during their relatively long lives (weeks to months) during which females experience multiple mating events and retain the eggs internally for long periods of time to allow egg maturation and host plant discovery (31). This ecological strategy is significantly different from saturniids (16), (19), (49), which depend entirely upon the resources acquired during the larval period. Adult saturniids possess reduced or non-functional mouthparts, lay eggs almost immediately after mating, do not have ears or "jamming" organs, and exhibit erratic evasive flight (31). 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.

To better answer questions regarding bombycoid evolution, we applied Anchored Hybrid Enrichment (AHE) targeted-sequencing phylogenomics (50), sampling 571 loci in 117 taxa across Bombycoidea. In order to do this, we developed a new AHE probe set ("BOM1"), redesigned from the "LEP1" probe set of Breinholt et al. (25), that more efficiently captures loci across the Bombycoidea, allowing us to establish a robust evolutionary framework for future research. This new probe set also includes loci that were not in the original "LEP1" (25) set, but were part of existing Sanger-sequencing datasets (11), (13), (24), (51). Finally, we present a modified DNA extraction protocol for use with historical natural history specimens, allowing researchers to tap into the biological wealth of data residing in the many museum collections around the globe.

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 provide mechanisms whereby different researchers can confidently and effectively use the same loci for independent projects, allowing for the combination of data across studies.

To build a more Bombycoidea-specific AHE probe set and phylogenomic dataset, we began with the Lepidoptera Agilent Custom SureSelect Target Enrichment "LEP1" probe kit, designed for 855 loci (21). We then modified this kit, henceforth the new "BOM1" kit, by evaluating which loci were most phylogenetically-informative within the superfamily, optimizing the set of probes to recover these loci, and including 24 previously-sequenced "legacy" Sanger-sequenced loci, e.g., CO1 for species confirmation in BOLD (1), from (11), (13), (24), and (51) (see Supp. Table 1 for loci names), as well as eight vision-related genes. The probes for these vision-related genes are based phototransduction genes mined from eye or head transcriptomes (unpublished; generated by AYK), and are included for future analyses to investigate their evolution across the group.

132 Phylogenetically-informative loci were determined by examining the sequence variation
133 present in 56 bombycoid species from across the taxonomic breadth of the superfamily (55 AHE
134 samples, pulled from (25) or generated for this study, and the *Bombyx mori* reference genome
135 (52)). These samples were sequenced and processed using the "LEP1" probe kit and Breinholt
136 et al. (25) bioinformatics pipeline (53). Individual gene trees were inferred for each of the 855
137 loci, using Maximum Likelihood in RAxML v8.2 (54), under a GTRGAMMA model of evolution
138 with 100 non-parametric bootstraps for node support. For each tree, phylogenetic
139 informativeness was calculated using PhyDesign (55), (56), (57),
140 http://phydesign.townsend.yale.edu. Parsimony informative characters and number of
141 segregating sites were calculated using the R packages 'phyloch' (58) and 'ape' (59)
142 respectively. Additionally, each phylogeny was scrutinized by eye to determine the level of
143 informative branching (i.e., branch lengths and topological patterns). Loci deemed to be
144 phylogenetically uninformative, or those that were capturing poorly (<60% of sampled species
145 represented for a locus), were removed from consideration. The final "BOM1" probe kit
146 comprised 571 loci, 539 of which came from the original "LEP1" kit.
147
148 *Taxon sampling*
149 To build a "backbone" phylogeny of the Bombycoidea, we sampled all major lineages
150 (i.e., families, subfamilies, and tribes), with the exception of three rare, species-poor groups:
151 subfamily Munychryiinae (Anthelidae) and tribes Sataspedini and Monardini (Sphingidae), for
152 which representative samples were unavailable for DNA sequencing. In total, 115 ingroup
153 Bombycoidea species from 97 genera were included in the phylogenetic analysis, as well as two
154 lasiocampid outgroups (Supp. Table 2).
155 Specimens were obtained from fieldwork, historically preserved dry collections (Supp
156 Table 1), and molecular tissue collections. Field-collected specimens were stored in ≥95%
157 ethanol, RNAlater, or papered and dried with silica gel, and stored in freezers at -80°C. Genomic
158 DNA was extracted using OmniPrep Genomic DNA Extraction Kits (G-Biosciences, St. Louis, MO,
159 USA) and Dneasy Blood and Tissue Kits (Qiagen, Valencia, CA, USA). DNA concentration was
160 evaluated through agarose gel electrophoresis and fluorometery using a Qubit 2.0 (Invitrogen,
161 Thermo Fisher Scientific, Carlsbad, CA, USA). Library preparation, hybridization enrichment, and
162 Illumina HiSeq 2500 sequencing (PE100) was carried out by RAPiD Genomics (Gainesville, FL,
163 USA). All specimen wing vouchers and tissue storage methods follow Cho et al. (60). All ethanol,
164 RNAlater, freshly papered specimens, or previously extracted DNA were stored at -80°C in the
165 molecular tissue collection of the MGCL.
166
167 *DNA extraction protocol for museum specimens*
168 We evaluated the efficacy of obtaining DNA from historical museum specimens because
169 there is great interest in understanding the feasibility of this approach (28), (61), (62), (63). The
170 Lepidoptera specimens evaluated herein, were "field-pinned" (never rehydrated), "papered"
171 (stored in an envelope and kept dry since collected, thus not rehydrated or pinned), and
172 "traditionally-pinned" specimens (dried, rehydrated, and subsequently pinned) – the
173 historically most common method of Lepidoptera specimen storage. Collecting dates ranged
174 from 1987 to 2017 (Supp. Table 2). The samples themselves were not preserved for molecular
175 work and information about potential contaminants and/or extraction inhibitors (i.e.

176 fumigation compounds, other chemicals) was unavailable. In some cases there was little soft-
177 tissue to extract from within the abdomens, having degraded over time. Being that these
178 samples were not kept in molecular-grade conditions, many were contaminated with fungal
179 and bacterial growth, which could be identified visually by the spores left on the bodies or by
180 the smell of decay. These factors along with the amount of sclerotized tissue present and the
181 fact that the abdomens used needed to stay intact (not homogenized) for dissecting purposes,
182 made extracting good quality genomic DNA challenging.
183 The method, detailed in the Supplemental information, attempts to account for several
184 factors: the amount of degraded tissue, the presence of eggs, the relative fat content, and the
185 overall abdomen size. Many of the commercial DNA extraction kits on the market (including the
186 Omni Prep kit used in this study) recommend using 10mg-20mg of well preserved, soft tissue
187 for the extraction process. Given that the museum specimens used had been desiccated for
188 many years, a number of abdomens had little to no visible internal soft tissue left. We increased
189 the amount of proteinase K to lysis buffer ratio in hopes that the proteinase K would digest the
190 remaining DNA rich material in solution. The relative fat content of samples is an issue because
191 lipid-rich tissue can interfere with the digestion of the soft tissue as well as change the
192 chemistry of the DNA isolation buffers. The lipid-rich layer in the Chloroform phase separation
193 step made pipetting of the DNA rich supernatant difficult as well. Any specimen that appeared
194 to be "greasy" or seemed to have an oily film on the abdomen was not used for extraction.
195 Finally, the overall size of the abdomen dictated the amount of lysis buffer needed in order to
196 sufficiently submerge the abdomen and reach the available soft tissue, which then dictated the
197 volume of reagents needed for the remainder of the DNA isolation process. Lastly, and
198 importantly, the modified protocol will allow a user not only to obtain DNA, but also easily
199 prepare the genitalia for taxonomic or future morphological work using standard techniques.
200 Detailed explanation of the methods can be found in the Supplemental information.
201
202 *Bioinformatics*
203 The bioinformatics pipeline of Breinholt et al. (25) was used to clean and assemble raw
204 Illumina reads for each AHE locus. The pipeline uses a probe-baited iterative assembly that
205 extends beyond the "probe" region, checks for quality and cross contamination due to barcode
206 leakage, removes paralogs, and returns a set of aligned orthologs for each locus and taxon of
207 interest. To accomplish these tasks, the pipeline uses the *Bombyx mori* genome (52), and
208 "LEP1" or "BOM1" reference libraries – depending on the kit used for sequencing.
209 Loci for phylogenetic analysis were selected by applying a cutoff of ≥40% sampled taxa
210 recovery – meaning that for a locus to be included in the analysis, the locus had to be recovered
211 in at least 40% of the sampled taxa. The pipeline then evaluates density and entropy at each
212 site of a nucleotide sequence alignment. We elected to trim with entropy and density cutoffs
213 only in flanking regions, allowing the "probe" region (exon) to be converted into amino acid
214 sequences. For a site (outside of the "probe" region) to remain, that site must pass a 60%
215 density and 1.5 entropy cutoff, rejecting sites that fail these requirements. A higher first value
216 (60) increases the coverage cutoff (e.g., a site is kept if 60% of all taxa are represented at that
217 site). A higher second value (1.5) increases the entropy cutoff (i.e., entropy values represent the
218 amount of saturation at a site); sites with values higher than 1.5 possess higher saturation and
219 are thus deleted). AliView v. 1.18 (64) was used to check for frame shifts, recognize and remove

220 stop codons (if present), and edit sequencing errors or lone/dubious indels. Reading frames
221 were manually checked after "probe" region extraction prior to translation. Translation was
222 carried out with FASconCAT-G v1.02 (65). Because flanking sequences are generally "non-
223 coding" and sites have been deemed homologous (see 25), these flanking sequence regions
224 before and after the "probe" regions (exons) were separated from the exons, combined, and
225 treated together as an independent partition.
226 Of the 115 bombycoid and two outgroup specimens, 110 were sequenced directly using
227 the AHE target capture sequencing, of which 68 were sequenced using the "BOM1" kit and 42
228 using the "LEP1" kit. Seven specimens had their AHE loci "probe" regions mined from previously
229 sequenced transcriptomes or the *Bombyx mori* genome (Supp. Table 2). These specimens did
230 not have "flanking" data because of their transcriptome-data origin. All specimens were
231 processed using either the 'Bmori' (for "LEP1") or 'BOM1' (for "BOM1") reference libraries.
232 Previous scripts are available in Dryad (53). For reproducibility, instructions on how to use the
233 pipeline and additional scripts not used by Breinholt et al. (25) are provided in Dryad (available
234 upon publication).
235
236 *Phylogenetics*
237 Lepidoptera AHE probe sets comprise highly-conserved coding "probe" regions (i.e.,
238 exons) and more variable, generally non-coding "flanking" regions (e.g., introns or intergenic
239 regions) located on either side of the "probe" region (25). We evaluated both nucleotide and
240 amino acid datasets to examine phylogenetic signal and the role that saturation may play in the
241 data (see 19). Using the Breinholt et al. (25) pipeline and scripts (53), also included herein, three
242 datasets were built for phylogeny inference: 1) AA = an amino acid supermatrix composed of
243 translated "probe" region loci; 2) Pr+Fl = a "probe" + "flanking" supermatrix; and 3) ASTRAL =
244 the individual loci from the Pr+Fl supermatrix, used for individual gene tree inference and then
245 species tree estimation. Additional analyses (i.e., "probe" region-only, as nucleotides) were
246 investigated, but are not reported here because their outcomes did not differ from those
247 reported.
248 Concatenated supermatrices were assembled using FASconCAT-G v1.02 (65).
249 Phylogenetic inference was performed in a maximum likelihood (ML) framework using IQ-TREE
250 MPI multicore v1.5.3 (66). For nucleotide datasets, the '–m TEST' command was used in IQ-
251 TREE to perform a model of evolution search resembling that of jModelTest/ProtTest (66),
252 applying models to each partition (AA) or site (Pr+Fl) to determine the best model of amino acid
253 or nucleotide substitution. For all inferences, we performed 1000 random addition sequence
254 (RAS) replicates, and 1000 replicates each for both ultrafast bootstraps (UFBS) ('–bb' command)
255 and SH-aLRT tests ('-alrt' command). The SH-like approximate likelihood ratio test (SH-aLRT)
256 estimates branch support values that have been shown to be as conservative as the commonly
257 used non-parametric bootstrap values (67). SH-aLRT and bootstrap values tend to agree for
258 data sets with strong phylogenetic signal (i.e., datasets with loci that are "sufficiently" large and
259 tips that share "sufficiently" diverged sequences). Disagreements in branch support are thought
260 to arise as a consequence of small sample size, insufficient data, or saturated divergence levels
261 (see 68). We classified nodes as "robust" if they were recovered with support values of UFBS ≥
262 95and SH-aLRT ≥ 80 (67), (68).

Because concatenation can be misleading when there are high levels of incomplete lineage sorting or deep coalescence (69), we assessed the impact of potential gene-tree discordance (70), (71), (72) 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 (73). 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 (74). To evaluate 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 (75). 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 as supplementary materials on Dryad (available upon publication).

Rogue taxon & outlier locus analyses

We investigated whether our molecular data included rogue taxa or outlier loci that were potentially influencing our phylogenetic inferences. A rogue taxon analysis was carried out using the online version of RogueNaRok (76), (77), 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) (78) in R (on the 650 gene trees used in the ASTRAL analysis). No rogue taxa or loci were found and therefore taxa were not pruned in 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. Instead of simply calculating species diversity per clade and assuming extant diversity is a true indicator of increases in diversification rate, an approach prone to biases, we applied BAMM (79) and 'BAMMtools' (80) to infer the number and location of macroevolutionary rate shifts across our phylogeny, as well as to visualize the 95% credible set of shift configurations. Because the Bombycoidea fossil record is poor (81), it is difficult to infer a robust dated phylogeny and therefore we chose not to do so for this study. We took the ML best tree and converted it into a relative-rate scaled ultrametric tree using the 'chronopl' command in the R package 'ape' (59). 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 and Rougiere 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' (82). The first 20% of runs were discarded as burn-in.

Results

While there is no guaranteed method for extracting high quality Lepidoptera gDNA from non-molecular grade specimens, our abdomen soaking approach proved relatively consistent and yielded us enough target DNA to proceed with AHE sequencing. The datasets built for phylogeny inference contained: 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.

Increasing the number of loci and species 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 lineages robustly supported, and all bombycoid families *sensu* Zwick (21) and Zwick et al. (13) 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 Supplemental), our preferred phylogeny was the "probe + flanking" (Pr+Fl) dataset. All family-level placements of genera *sensu* Kitching and Rougiere et al. (1) were supported, with the exception of *Arotros* Schaus, a genus long considered to be an epiine bombycid (83). *Arotros* is clearly nested within Apatelodidae in our tree (Fig. 1; Supp. Fig. 1) and is hereby transferred to Apatelodidae (see supplemental text for additional details on this transfer). The family-level rearrangements of Zwick (21) 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 a monophyletic group.

We find broad congruence with the major groupings designated by Zwick (21), though internal relationships within these groups may not match previous phylogenies exactly. Historically, the most problematic familial placement in the superfamily has been the Bombycidae *sensu stricto*. Traditional Sanger-sequencing phylogenetics based on a handful of gene regions, applied in (13), (18), and (21), 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) clade, as seen in both the AA and Pr+Fl datasets, but not in the ASTRAL inference – an outcome that mirrors traditional studies that used low numbers of loci and lacked the phylogenetic signal to confidently resolve relationships among these three families (Supp. Trees 1-3). The 'CAPOPEM' group (Carthaeidae, Anthelidae, Phiditiidae, and Endromidae) is recovered in all three analyses, though in the ASTRAL inference this group is nested within the 'SBS' group (Sphingidae, Bombycidae and Saturniidae) - due to placement of the Bombycidae (Supp. Fig. 2). The relative interfamilial relationships within the 'CAPOPEM' clade were not robustly supported in Zwick et al. (13), and although our more refined taxon sampling and increased genetic coverage confidently solidifies the group, the data still fail to fully resolve these relationships across analyses (Supp. Fig. 2). 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). This discordance is important to note considering the disjunct geographic distribution of these families. Among these families, only Phiditiidae is found in (and is endemic to) the New World, with each of the other families being restricted to the Old World. Intrafamilial relationships in the 'SBS' and 'CAPOPEM' groups change depending on which of our datasets are used to infer the phylogeny (Fig. 1, Supp. Figs. 1-3, Supp. Trees 1-3). Lastly, the 'BALE' group (Brahmaeidae, Apatelodidae, and Eupterotidae) is recovered in all three analyses.

When we investigate intrafamilial relationships, the Bombycidae (disregarding the taxonomically "incorrect" *Arotros*, mentioned above; see Supplemental text) possesses distinct and monophyletic subfamilies Bombycinae (Old World) and Epiinae (New World; Fig. 1; Supp. Fig. 1). Within the Eupterotidae, the subfamilies Striphnopteryginae and Janinae are monophyletic, but others are not (Fig. 1; Supp. Fig. 1). Within the Saturniidae, the Pr+Fl and AA topologies are similar to Regier et al. (51) (Fig. 1 & Supp. Fig. 3). 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 subfamily relationship difference lies with the Agliinae, sister to the Salassinae + Saturniinae (Pr+Fl) or sister to the Arsenurinae (AA, also (13) and (51)). 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 (Supp. Fig. 3), though with low support. Within the Sphingidae, our topologies (both Pr+Fl & AA) are quite different from Zwick et al. (13). Instead of the Smerinthinae being paraphyletic, the Smerinthinae are monophyletic and the Sphinginae are polyphyletic (Fig. 1; Supp. Fig. 1). This outcome is robustly supported in the Pr+Fl and ASTRAL phylogenies, though internal structuring is slightly different. In the AA phylogeny, while the placement of the Macroglossinae, Langiinae and Sphinginae are robustly supported (and the same topology as in Pr+Fl), low internal support obfuscates relationships within the Smerinthinae.

Importantly, a number of subfamilies and genera are not monophyletic, based on our sampling and the most current taxonomic definition (see Kitching and Rougerie et al. (1)) (Fig. 1; Supp. Fig. 1). 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, the genus *Anthela* is paraphyletic due to the placement of *Nataxa* and *Pterolocera*, a finding in agreement with (21) . And lastly, within the Sphingidae, Sphinginae is polyphyletic due to the placement of *Pentateucha* as sister to the Langiinae, and the Smerinthinae genus *Polyptychus* is paraphyletic.

To quantify whether certain bombycoid lineages exhibit differing diversification rates, we evaluated how the interplay between speciation and extinction has changed over relative time rather than using just the simple sums of described species in each family. At the family scale, a major diversification rate shift occurs along the lineage leading to the Saturniidae and Sphingidae (Fig. 1). Unfortunately, due to the poor Bombycoidea fossil record and the poor ability of computational approaches to estimate extinction rates, we are not able to discern

395 whether the differences that exist are due to increases or decreases in speciation, although the
396 raw numbers, based on (1), do agree with the outcomes found herein: Sphingidae (1,602 spp.);
397 Saturniidae (3,454 spp.); Bombycidae (202 spp.); Endromidae (70 spp.); Phiditiidae (23 spp.);
398 Carthaeidae (1 sp.); Anthelidae (94 spp.); Eupterotidae (396 spp.); Brahmaeidae (68 spp.);
399 Apatelotidae (182 spp.).
400
401 **Discussion**
402 Within the past decade, only a handful of studies have attempted to use molecular
403 sequence data to understand the phylogenetic relationships of the Bombycoidea (11), (13), (21)
404 and none corroborated the earlier subjective morphology-based hypotheses (20), (23). So, to
405 establish a foundational backbone for future evolutionary, comparative, and taxonomic studies,
406 we sampled exemplars from all major lineages in the superfamily and used Anchored Hybrid
407 Enrichment (AHE) phylogenomics to provide a robust phylogeny of the superfamily based on
408 the largest taxonomic and molecular sampling to date.
409 To achieve dense taxonomic coverage, we included many samples from natural history
410 collections, and modified previous DNA extraction methodology to obtain DNA for high-
411 throughput sequencing. We also redesigned the AHE probe set developed by Breinholt et al.
412 (25) and modified it to more efficiently recover phylogenetically-informative loci within the
413 Bombycoidea, as well as adding "legacy" Sanger-sequenced loci for increasing the usability of
414 the data (e.g., integration of "legacy" datasets and CO1 for species identification). To assist
415 future usage of this and other AHE probe sets, we expanded the command log instructions of
416 Breinholt et al. (25) to include step-by-step instructions for running the bioinformatics pipeline.
417 Our phylogeny largely reinforces the phylogenetic results of earlier molecular works,
418 although with topological differences within families (i.e., some subfamilies and genera are not
419 recovered as monophyletic in our study). Topological discordances with trees published in
420 previous works are likely to be the product of increased locus sampling, which provided
421 significantly more phylogenetic information. For example, our results provide a well-supported
422 placement of the historically problematic family Bombycidae *sensu stricto*, as the sister lineage
423 to Saturniidae + Sphingidae. An earlier study (19) had provided some evidence in support of this
424 relationship, but its sampling was very limited. Additionally, it is quite likely that for the deep to
425 mid-timescale nodes, the AA phylogeny is the best hypothesis for Bombycoidea relationships.
426 Our study highlights how morphological convergence in Bombycoidea has confused our
427 understanding of their evolution. This convergence is sometimes exceptional. For example, the
428 bombycid genus *Rotunda*, endemic to the Old World, and the Apatelodidae *Arotros*, endemic to
429 the New World, are widely separated on the phylogenetic tree (Fig. 1; Supp. Fig. 1) and
430 geographically, yet are astonishingly similar in body size, wing shape, and phenotypic
431 appearance (see Systematics section in Supplemental Information). Additionally, when we
432 examine the 'CAPOPEM' group closely, we see a biogeographical distribution that, based on the
433 AA and ASTRAL topologies, appears to be a more biologically realistic and parsimonious
434 evolutionary history due to the placement of the Neotropical Phiditiidae as sister to the
435 remaining and Old World 'CAPOPEM' families (Supp. Fig. 2).
436 One of the most surprising results of the present study is the dramatic shift in
437 diversification rates within both the Sphingidae and Saturniidae compared to other bombycoid
438 lineages. These differences in "survivability" may be explained by their very different life-

history strategies, morphology (e.g., body sizes and wing shapes), flight speed and maneuverability (31), as well as very different anti-bat strategies (30), (32), (33), (47). Although we did not conduct a divergence time estimation analysis in this study, the origin of the 'SBS' group has been postulated to be around 50 mya (85), with hawkmoths originating soon thereafter (33). 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 is in part due to bat-related selection pressures resulting in diverse anti-bat traits, traits that other bombycoid lineages lack.

However, the hawkmoth and silk moth evolutionary story is certainly more complicated than simply reflecting their interactions with bats. As evolution is often complex, there are undoubtedly a suite of factors playing significant roles in their diversity of form and function. Although we hypothesize 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.

Acknowledgements

We thank our collaborators Seth Bybee, Deborah Glass, Yash Sondhi, Jamie Theobald, and Emmanuel Touissant. We are indebted to those individuals that provided specimens for this project, contributed help in identifying specimens, or in other ways: Samantha Epstein, John Heppner, Geena Hill, Nick Homziak, Jacqueline Miller, Charles Mitter, Kim Mitter, Lary Reeves, Andrei Sourakov, and Andrew Warren.

Funding

This work was partly supported by NSF grants [IOS-1121807, JRB], [IOS-1121739, AYK], [DBI-1349345, AYK], [DEB-1557007, AYK], and [PRFB-1612862, CAH]; a NERC grant [NE/P003915/1, IJK]; National Geographic Society grants [CRE 9944-16, JRB]; and postdoctoral support from the Florida Museum of Natural History (to CAH).

Author Contributions

CAH, AYK, JWB, IJK, MT, AZ designed the study. CAH and JWB analyzed data. CAH, AYK, RAS wrote the paper. CAH, AYK, JRB, KD, IJK, RAS, AZ collected, identified, and processed specimens. All authors discussed results and contributed to the writing of the manuscript.

Supplemental Information

The datasets supporting this article have been uploaded as part of the supplementary material or can be found in Dryad, (available upon publication). Supplemental Information includes details of the DNA extraction protocol using museum specimens, three figures, one table, detailed systematic information on the Bombycoidea families, bioinformatics pipeline scripts, detailed instructions on running the pipeline, all alignment FASTA files (nucleotides, amino acids, supermatrices, and individual loci), partition files, tree files, and other essential

483 data files used for phylogenetic inference and analysis. Previous Breinholt et al. bioinformatics
484 scripts are available in Dryad (53).
485
486 **Figures**
487 Figure 1. Maximum likelihood tree of Bombycoidea, based on 650 AHE loci. All nodes are
488 supported by ≥95% bootstrap values unless otherwise noted. Branch color indicates the
489 estimated diversification rate, with warmer colors representing lineages with higher rates.
490 Major taxonomic groups such as families and subfamilies are labeled. Photographs represent
491 lineages sampled in the phylogeny. Species-level diversity, based on Kitching and Rougerie et al.
492 (1), are also noted next to Family.
493
494 **References**

495 1. Kitching I, Rougerie R, Zwick A, St Laurent R, Ballesteros Mejia L, Kawahara A, et al. A
496 global checklist of the Bombycoidea (Insecta: Lepidoptera). BDJ. 2018 Feb
497 12;6(1):e22236–13.

498 2. Roe AD, Weller SJ, Baixeras J, Brown J, Cummings MP, Davis DR, et al. Evolutionary
499 framework for Lepidoptera model systems. In: Goldsmith M, Marec F, editors. Molecular
500 Biology and Genetics of the Lepidoptera. 2010. pp. 1–24.

501 3. Gregory D. Hawkmoth pollination in the genus *Oenothera*. Aliso A Journal of Systematic
502 and Evolutionary Botany. 1963;5(3):357–84.

503 4. Haber WA, Frankie GW. A Tropical Hawkmoth Community- Costa Rican Dry Forest
504 Sphingidae. Biotropica. 1989 Dec 9;21(2):155–72.

505 5. Nilsson LA, Jonsson L, Rason L, Randrianjohany E. Monophily and pollination mechanisms
506 in *Angraecum arachnites* Schltr. (Orchidaceae) in a guild of long-tongued hawk-moths
507 (Sphingidae) in Madagascar. Biological Journal of the Linnean Society. 1985;26(1):1–19.

508 6. Nilsson LA, Rabakonandrianina E, Pettersson B. Exact tracking of pollen transfer and
509 mating in plants. Nature. 1992;360(6405):666–8.

510 7. Wasserthal LT. The pollinators of the Malagasy star orchids *Angraecum sesquipedale*, *A.*
511 *sororium* and *A. compactum* and the evolution of extremely long spurs by pollinator shift.
512 Botanica Acta. 1997 Oct;110(5):343–59.

513 8. Nilsson LA. Deep flowers for long tongues. Trends In Ecology & Evolution. 1998
514 Jul;13(7):259–60.

515 9. Raguso RA, Willis MA. Synergy between visual and olfactory cues in nectar feeding by
516 naïve hawkmoths, *Manduca sexta*. Animal Behaviour. 2002 Nov 1;64(5):685–95.

517 10. Beck J, Kitching IJ, Linsenmair KE. Effects of habitat disturbance can be subtle yet
518 significant: biodiversity of hawkmoth-assemblages (Lepidoptera: Sphingidae) in

519 Southeast-Asia. Biodiversity and Conservation. 2nd ed. Kluwer Academic Publishers;
520 2006 Jan;15(1):465–86.

521 11. Regier JC, Cook CP, Mitter C, Hussey A. A phylogenetic study of the “bombycoid
522 complex” (Lepidoptera) using five protein-coding nuclear genes, with comments on the
523 problem of macrolepidopteran phylogeny. Systematic Entomology. 2008;33:175–89.

524 12. Meusemann K, Reumont von BM, Simon S, Roeding F, Strauss S, Kück P, et al. A
525 phylogenomic approach to resolve the Arthropod Tree of Life. Mol Biol Evol. 2010 Oct
526 15;27(11):2451–64.

527 13. Zwick A, Regier JC, Mitter C, Cummings MP. Increased gene sampling yields robust
528 support for higher-level clades within Bombycoidea (Lepidoptera). Systematic
529 Entomology. 2011;36(1):31–43.

530 14. Kim MJ, Kang AR, Jeong HC, Kim K-G, Kim I. Reconstructing intraordinal relationships in
531 Lepidoptera using mitochondrial genome data with the description of two newly
532 sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera:
533 Lycaenidae). Mol Phylogenet Evol. Elsevier Inc; 2011 Nov 1;61(2):436–45.

534 15. Simon S, Narechania A, Desalle R, Hadrys H. Insect phylogenomics: exploring the source
535 of incongruence using new transcriptomic data. Genome Biol Evol. Oxford University
536 Press; 2012 Dec 1;4(12):1295–309.

537 16. Regier JC, Mitter C, Zwick A, Bazinet AL, Cummings MP, Kawahara AY, et al. A large-scale,
538 higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and
539 butterflies). Moreau CS, editor. PLoS ONE. 2013 Mar 12;8(3):e58568–23.

540 17. Timmermans MJTN, Lees DC, Simonsen TJ. Towards a mitogenomic phylogeny of
541 Lepidoptera. Mol Phylogenet Evol. Elsevier Inc; 2014 Oct 1;79(C):169–78.

542 18. Wang X, Chen Z-M, Gu X-S, Wang M, Huang G-H, Zwick A. Phylogenetic relationships
543 among Bombycidae s.l. (Lepidoptera) based on analyses of complete mitochondrial
544 genomes. Systematic Entomology. John Wiley & Sons, Ltd; 2018 Nov 20;20:537–9.

545 19. Breinholt JW, Kawahara AY. Phylotranscriptomics: saturated third codon positions
546 radically influence the estimation of trees based on next-gen data. Genome Biol Evol.
547 2013 Nov 1;5(11):2082–92.

548 20. Minet J. The Bombycoidea: Phylogeny and higher classification (Lepidoptera, Glossata).
549 Entomologica Scandinavica. 1994;25(1):63–88.

550 21. Zwick A. Molecular phylogeny of Anthelidae and other bombycoid taxa (Lepidoptera:
551 Bombycoidea). Systematic Entomology. 2008 Jan 7;33:190–209.

552 22. Minet J. Tentative reconstruction of the ditrysian phylogeny (Lepidoptera, Glossata).
553 Entomologica Scandinavica. 1991;22(1):69–95.

554 23. Lemaire C, Minet J. 18. The Bombycoidea and their relatives. In: Kristensen NP, editor.
555 Lepidoptera, Moths and Butterflies vol 1 Evolution, Systematics, and Biogeography
556 Handbook of Zoology. Walter de Gruyter, Berlin; 1998. pp. 322–53.

557 24. Regier JC, Zwick A, Cummings MP, Kawahara AY, Cho S, Weller S, et al. Toward
558 reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia):
559 an initial molecular study. BMC Evolutionary Biology. BioMed Central; 2009;9(1):280–21.

560 25. Breinholt JW, Earl C, Lemmon AR, Lemmon EM, Xiao L, Kawahara AY. Resolving
561 relationships among the megadiverse butterflies and moths with a novel pipeline for
562 anchored phylogenomics. Syst Biol. 2018;67(1):78–93.

563 26. Espeland M, Breinholt J, Willmott KR, Warren AD, Vila R, Toussaint EFA, et al. A
564 comprehensive and dated phylogenomic analysis of butterflies. CURBIO. Elsevier Ltd;
565 2018 Mar 5;28(5):770–5.

566 27. Johns CA, Toussaint EFA, Breinholt JW, Kawahara AY. Origin and macroevolution of
567 micro-moths on sunken Hawaiian Islands. Proc Biol Sci. The Royal Society; 2018 Aug
568 29;285(1885):20181047–10.

569 28. St Laurent RA, Hamilton CA, Kawahara AY. Museum specimens provide phylogenomic
570 data to resolve relationships of sack-bearer moths (Lepidoptera, Mimallonoidea,
571 Mimallonidae). Systematic Entomology. Wiley/Blackwell (10.1111); 2018 May
572 31;43(4):729–61.

573 29. Toussaint EFA, Breinholt JW, Earl C, Warren AD, Brower AVZ, Yago M, et al. Anchored
574 phylogenomics illuminates the skipper butterfly tree of life. BMC Evolutionary Biology.
575 BMC Evolutionary Biology; 2018 Jun 16;18(101):1–11.

576 30. Rubin JJ, Hamilton CA, McClure CJW, Chadwell BA, Kawahara AY, Barber JR. The
577 evolution of anti-bat sensory illusions in moths. Science Advances. American Association
578 for the Advancement of Science; 2018 Jul;4(7).

579 31. Janzen D. Two ways to be a tropical big moth: Santa Rosa saturniids and sphingids.
580 Oxford surveys in evolutionary biology. 1984;1(85):140.

581 32. Barber JR, Kawahara AY. Hawkmoths produce anti-bat ultrasound. Biology Lett. 2013
582 May 8;9(4):20130161–1.

583 33. Kawahara AY, Barber JR. Tempo and mode of antibat ultrasound production and sonar
584 jamming in the diverse hawkmoth radiation. Proceedings of the National Academy of
585 Sciences. 2015 May 19;112(20):6407–12.

586 34. Kawahara AY, Plotkin D, Hamilton CA, Gough H, Laurent RS, Owens HL, et al. Diel
587 behavior in moths and butterflies: a synthesis of data illuminates the evolution of
588 temporal activity. Organisms Diversity & Evolution; 2017 Dec 4;:1–15.

589 35. Jones KE, Bininda-Emonds ORP, Gittleman JL. Bats, clocks, and rocks: diversification
590 patterns in Chiroptera. Evolution. 2005 Oct;59(10):2243–55.

591 36. Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ. A molecular
592 phylogeny for bats illuminates biogeography and the fossil record. Science. American
593 Association for the Advancement of Science; 2005 Jan 28;307(5709):580–4.

594 37. Miller-Butterworth CM, Murphy WJ, O'Brien SJ, Jacobs DS, Springer MS, Teeling EC. A
595 family matter: conclusive resolution of the taxonomic position of the long-fingered bats,
596 *Miniopterus*. Mol Biol Evol. 2007 Jul;24(7):1553–61.

597 38. Shi JJ, Rabosky DL. Speciation dynamics during the global radiation of extant bats.
598 Evolution. 2015 Jun;69(6):1528–45.

599 39. Lei M, Dong D. Phylogenomic analyses of bat subordinal relationships based on
600 transcriptome data. Nature Publishing Group. 2016 Jun 13;6(1):27726.

601 40. Roeder KD, Treat AE, Vandeberg JS. Auditory sense in certain hawkmoths. Science.
602 1968;159:331–3.

603 41. Spangler HG. Moth hearing, defense, and communication. Annual Review of Entomology.
604 Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA;
605 1988;33(1):59–81.

606 42. Göpfert MC, Surlykke A, Wasserthal LT. Tympanal and atympanal “mouth-ears” in
607 hawkmoths (Sphingidae). P Roy Soc B-Biol Sci. The Royal Society; 2002 Jan
608 7;269(1486):89–95.

609 43. Hofstede ter HM, Goerlitz HR, Ratcliffe JM, Holderied MW, Surlykke A. The simple ears of
610 noctuoid moths are tuned to the calls of their sympatric bat community. J Exp Biol. The
611 Company of Biologists Ltd; 2013 Nov 1;216(Pt 21):3954–62.

612 44. Barber JR, Conner WE. Acoustic mimicry in a predator-prey interaction. P Natl Acad Sci
613 Usa. National Academy of Sciences; 2007;104(22):9331–4.

614 45. Lamarre GPA, Mendoza I, Rougerie R, Decaëns T, Hérault B, Bénéluz F. Stay out (almost)
615 all night: contrasting responses in flight activity among tropical moth assemblages.
616 Neotrop Entomol. 2015 Feb 10;44(2):109–15.

617 46. Zhong M, Hill GM, Gomez JP, the DPTJO, 2016. Quantifying wing shape and size of
618 saturniid moths with geometric morphometrics. Journal of the Lepidopterists' Society.
619 2016;70(2):1–9.

620 47. Barber JR, Leavell BC, Keener AL, Breinholt JW, Chadwell BA, McClure CJW, et al. Moth
621 tails divert bat attack: evolution of acoustic deflection. Proceedings of the National
622 Academy of Sciences. National Acad Sciences; 2015 Mar 3;112(9):2812–6.

623 48. Tammaru T, Haukioja E. Capital breeders and income breeders among Lepidoptera:
624 consequences to population dynamics. Oikos. 1996;77(3):561–4.

625 49. Kawahara AY, Breinholt JW. Phylogenomics provides strong evidence for relationships of
626 butterflies and moths. P Roy Soc B-Biol Sci. 2014 Jun 18;281(1788):20140970–0.

627 50. Lemmon AR, Emme SA, Lemmon EM. Anchored Hybrid Enrichment for massively high-
628 throughput phylogenomics. Syst Biol. 2012 Aug 11;61(5):727–44.

629 51. Regier JC, Grant MC, Mitter C, Cook CP, Peigler RS, Rougerie R. Phylogenetic relationships
630 of wild silkmoths (Lepidoptera: Saturniidae) inferred from four protein-coding nuclear
631 genes. Systematic Entomology. 2008 Feb 4;33(2):219–28.

632 52. Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Bin Li, et al. A draft sequence for the genome of the
633 domesticated silkworm (*Bombyx mori*). Science. American Association for the
634 Advancement of Science; 2004 Dec 10;306(5703):1937–40.

635 53. Breinholt JW, Earl C, Lemmon AR, Lemmon EM, Xiao L, Kawahara AY. Data from:
636 Resolving relationships among the megadiverse butterflies and moths with a novel
637 pipeline for Anchored Phylogenomics. Dryad Digital Repository; 2017.

638 54. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
639 phylogenies. Bioinformatics. 2014.

640 55. Mayrose I, Graur D, Ben-Tal N, Pupko T. Comparison of site-specific rate-inference
641 methods for protein sequences: empirical Bayesian methods are superior. Mol Biol Evol.
642 2004 May 21;21(9):1781–91.

643 56. Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. Springer,
644 New York, NY; 2005. pp. 125–81.

645 57. López-Giráldez F, Townsend JP. PhyDesign: an online application for profiling
646 phylogenetic informativeness. BMC Evolutionary Biology. BioMed Central Ltd; 2011 May
647 31;11(1):152.

648 58. Heibl C. PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic
649 software packages. 2013.

650 59. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R
651 language. Bioinformatics. 2004 Jan 20;20(2):289–90.

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

655 61. Blaimer BB, Lloyd MW, Guillory WX, Brady SG. Sequence capture and phylogenetic utility
656 of genomic Ultraconserved Elements obtained from pinned insect specimens. O'Grady P,
657 editor. PLoS ONE. Public Library of Science; 2016 Aug 24;11(8):e0161531–20.

658 62. Mikheyev AS, Zwick A, Magrath MJL, Grau ML, Qiu L, Su YN, et al. Museum genomics
659 confirms that the Lord Howe Island stick insect survived extinction. CURBIO. Elsevier Ltd;
660 2017 Oct 23;27(20):3157–3161.e4.

661 63. Ruane S, Austin CC. Phylogenomics using formalin-fixed and 100+ year-old intractable
662 natural history specimens. Molecular Ecology Resources. Wiley/Blackwell (10.1111);
663 2017 Feb 13;17(5):1003–8.

664 64. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets.
665 Bioinformatics. 2014.

666 65. Kück P, Longo GC. FASconCAT-G: extensive functions for multiple sequence alignment
667 preparations concerning phylogenetic studies. Frontiers in Zoology. 2014;11(81).

668 66. Nguyen L-T, Schmidt HA, Haeseler von A, Minh BQ. IQ-TREE: a fast and effective
669 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015
670 Jan;32(1):268–74.

671 67. Minh BQ, Nguyen MAT, Haeseler von A. Ultrafast approximation for phylogenetic
672 bootstrap. Mol Biol Evol. 2013 Apr 15;30(5):1188–95.

673 68. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms
674 and methods to estimate Maximum-Likelihood phylogenies: assessing the performance
675 of PhyML 3.0. Syst Biol. 2010 May;59(3):307–21.

676 69. Mendes FK, Hahn MW. Why concatenation fails near the anomaly zone. Syst Biol. 2017
677 Jul 11;67(1):158–69.

678 70. Maddison W. Gene trees in species trees. Syst Biol. 1997 Sep 1;46(3):523–36.

679 71. Slowinski J, Page R. How should species phylogenies be inferred from sequence data?
680 Syst Biol. 1999;48(4):814–25.

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

683 73. Zhang C, Sayyari E, Mirarab S. ASTRAL-III: increased scalability and impacts of contracting
684 low support branches. In: RECOMB International Workshop on Comparative Genomics.
685 2017. pp. 53–75.

686 74. Mirarab S, Warnow T. ASTRAL-II: coalescent-based species tree estimation with many
687 hundreds of taxa and thousands of genes. Bioinformatics. 2015 Jun 9;31:i44–i52.

688 75. Sayyari E, Mirarab S. Anchoring quartet-based phylogenetic distances and applications to
689 species tree reconstruction. BMC Genomics. BioMed Central; 2016 Nov 11;17(10):101–
690 13.

691 76. Aberer AJ, Stamatakis A. A simple and accurate method for rogue taxon identification.
692 2011. pp. 118–22.

693 77. Aberer AJ, Krompass D, Stamatakis A. Pruning rogue taxa improves phylogenetic
694 accuracy: an efficient algorithm and webservice. Syst Biol. 2012 Nov 8;62(1):162–6.

695 78. de Vienne DM, Ollier S, Aguileta G. Phylo-MCOA: a fast and efficient method to detect
696 outlier genes and species in phylogenomics using multiple co-inertia analysis. Mol Biol
697 Evol. 2012 May 15;29(6):1587–98.

698 79. Rabosky DL. Automatic detection of key innovations, rate shifts, and diversity-
699 dependence on phylogenetic trees. Kolokotronis S-O, editor. PLoS ONE. 2014 Feb
700 26;9(2):e89543.

701 80. Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, et al. BAMMtools: an R
702 package for the analysis of evolutionary dynamics on phylogenetic trees. Kembel S,
703 editor. Methods in Ecology and Evolution. 2014 May 10;5(7):701–7.

704 81. Sohn JC, Labandeira CC, Davis DR, Mitter C. An annotated catalog of fossil and subfossil
705 Lepidoptera (Insecta: Holometabola) of the world. Zootaxa. 2012.

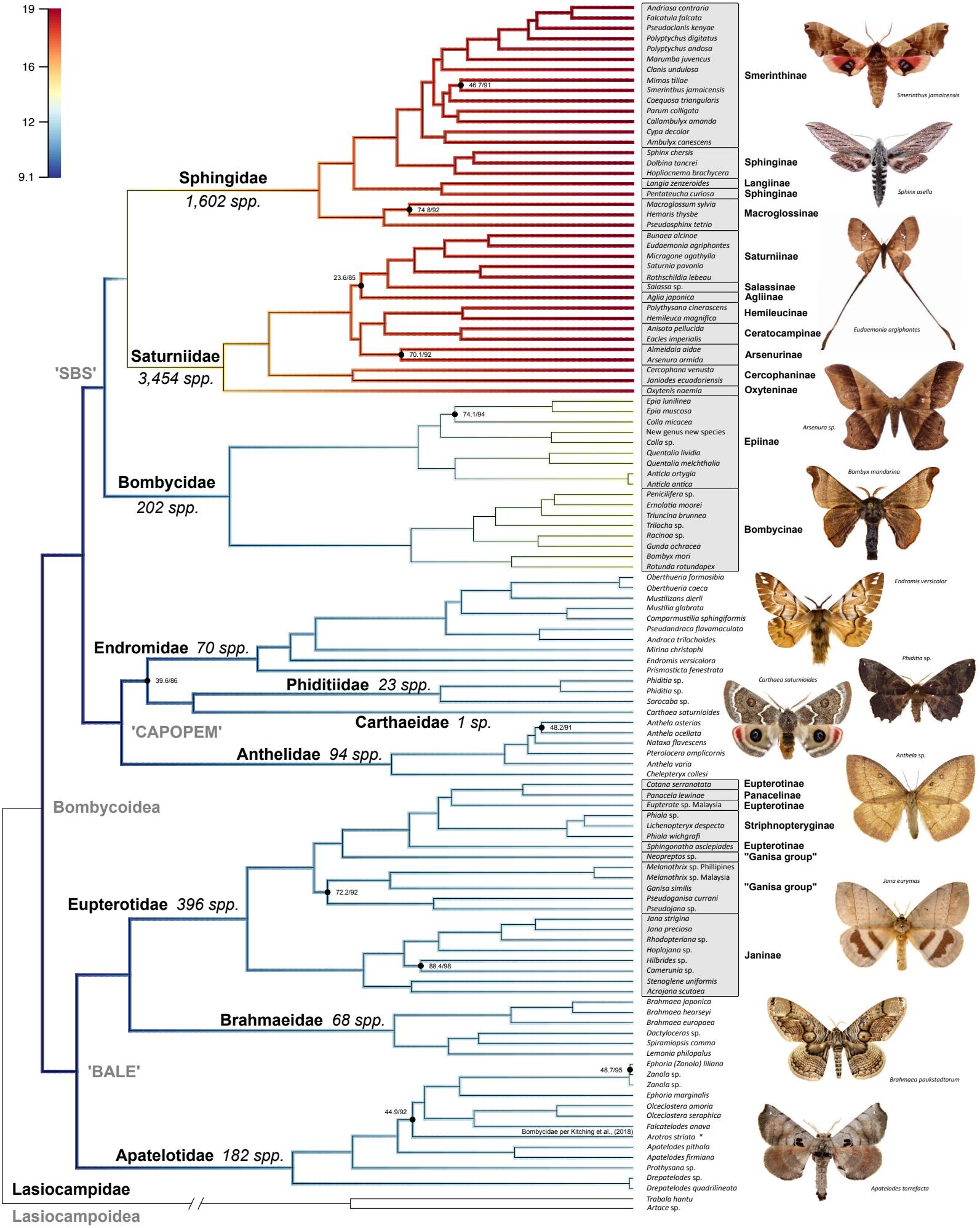
706 82. Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis
707 for MCMC. R News. 2006;6(1):7–11.

708 83. Schaus W. Familie Bombycidae (Epiinae). In: Seitz A, editor. Die Gross-Schmetterlinge der
709 Erde. Kernen, Stuttgart; 1929. pp. 675–711. (Die amerikanischen Spinner und
710 Schwärmer).

711 84. Nässig WA, Oberprieler RG. An annotated catalogue of the genera of Eupterotidae
712 (Insecta, Lepidoptera, Bombycoidea). Senckenbergiana biologica. 2008;88(1):53–80.

713 85. Misof B, Liu S, Meusemann K, Peters RS, Donath A. Phylogenomics resolves the timing
714 and pattern of insect evolution. Science. 2014;346(6210):763–7.

715


19
16
12
9.1
Sphingidae
1,602 spp.
Saturniidae
3,454 spp.
Bombycidae
202 spp.
Endromidae 70 spp.
Phiditiidae 23 spp.
Carthaeidae 1 sp.
Anthelidae 94 spp.
Eupterotidae 396 spp.
Brahmaeidae 68 spp.
Apatelotidae 182 spp.
Lasiocampidae
'SBS'
'CAPOPEM'
'BALE'
Bombycoidea
Lasiocampoidea
46.7/91
74.8/92
23.6/85
70.1/92
74.1/94
39.6/86
48.2/91
72.2/92
88.4/98
48.7/95
44.9/92
Bombycidae per Kitching et al., (2018)
Andriasa contraria
Falcatula falcata
Pseudoclanis kenyae
Polyptychus digitatus
Polyptychus andosa
Marumba juvencus
Clanis undulosa
Mimas tiliae
Smerinthus jamaicensis
Coequosa triangularis
Parum colligata
Callambulyx amanda
Cypa decolor
Ambulyx canescens
Sphinx chersis
Dolbina tancrei
Hopliocnema brachycera
Langia zenzeroides
Pentateucha curiosa
Macroglossum sylvia
Hemaris thysbe
Pseudosphinx tetrio
Bunaea alcinoe
Eudaemonia agriphontes
Micragone agathylla
Saturnia pavonia
Rothschildia lebeau
Salassa sp.
Aglia japonica
Polythysana cinerascens
Hemileuca magnifica
Anisota pellucida
Eacles imperialis
Almeidaia aidae
Arsenura armida
Cercophana venusta
Janiodes ecuadoriensis
Oxytenis naemia
Epia lunilinea
Epia muscosa
Colla micacea
New genus new species
Colla sp.
Quentalia lividia
Quentalia melchthalia
Anticla ortygia
Anticla antica
Penicilifera sp.
Ernolatia moorei
Triuncina brunnea
Trilocha sp.
Racinoa sp.
Gunda ochracea
Bombyx mori
Rotunda rotundapex
Oberthueria formosibia
Oberthueria caeca
Mustilizans dierli
Mustilia glabrata
Comparmustilia sphingiformis
Pseudandraca flavamaculata
Andraca trilochoides
Mirina christophi
Endromis versicolora
Prismosticta fenestrata
Phiditia sp.
Phiditia sp.
Sorocaba sp.
Carthaea saturnioides
Anthela asterias
Anthela ocellata
Nataxa flavescens
Pterolocera amplicornis
Anthela varia
Chelepteryx collesi
Cotana serranotata
Panacela lewinae
Eupterote sp. Malaysia
Phiala sp.
Lichenopteryx despecta
Phiala wichgrafi
Sphingonatha asclepiades
Neopreptos sp.
Melanothrix sp. Phillipines
Melanothrix sp. Malaysia
Ganisa similis
Pseudoganisa currani
Pseudojana sp.
Jana strigina
Jana preciosa
Rhodopteriana sp.
Hoplojana sp.
Hilbrides sp.
Camerunia sp.
Stenoglene uniformis
Acrojana scutaea
Brahmaea japonica
Brahmaea hearseyi
Brahmaea europaea
Dactyloceras sp.
Spiramiopsis comma
Lemonia philopalus
Ephoria (Zanola) liliana
Zanola sp.
Zanola sp.
Ephoria marginalis
Olceclostera amoria
Olceclostera seraphica
Falcatelodes anava
Arotros striata *
Apatelodes pithala
Apatelodes firmiana
Prothysana sp.
Drepatelodes sp.
Drepatelodes quadrilineata
Trabala hantu
Artace sp.
Smerinthinae
Sphinginae
Langiinae
Sphinginae
Macroglossinae
Saturniinae
Salassinae
Agliinae
Hemileucinae
Ceratocampinae
Arsenurinae
Cercophaninae
Oxyteninae
Epiinae
Bombycinae
Eupterotinae
Panacelinae
Eupterotinae
Striphnopteryginae
Eupterotinae
"Ganisa group"
"Ganisa group"
Janinae
Smerinthus jamaicensis
Sphinx asella
Eudaemonia argiphontes
Arsenura sp.
Bombyx mandarina
Endromis versicolor
Phiditia sp.
Carthaea saturnioides
Anthela sp.
Jana eurymas
Brahmaea paukstadtorum
Apatelodes torrefacta