**Supplementary File 3**

**Multispecies coalescent**

The multispecies coalescent, also known as gene tree-species tree methods, have proven to be a useful alternative to concatenation-based approaches, allowing estimation of phylogeny while accounting for different gene histories ([Brandley et al. 2015](#_ENREF_2); [Eytan et al. 2015](#_ENREF_5); [Peloso et al. 2016](#_ENREF_18); [Prum et al. 2015](#_ENREF_19); [Pyron et al. 2014](#_ENREF_20); [Ruane et al. 2015](#_ENREF_24); [Song et al. 2012](#_ENREF_26)). However, some have contended that the multispecies coalescent approach can in some cases produce a biologically unrealistic tree, or a tree with less resolution than ML trees based on concatenated data ([Kimball et al. 2013](#_ENREF_12); [McCormack et al. 2013](#_ENREF_14); [Mirarab et al. 2014](#_ENREF_16)). The use of multispecies coalescent for datasets based on transcriptomes, UCEs, and AHE data has been proposed to be problematic, as it has been suggested that these data types violate the multispecies coalescent model and the quality of these data may lead to incorrect species trees ([Gatesy and Springer 2014](#_ENREF_6); [Simmons and Gatesy 2015](#_ENREF_25); [Springer and Gatesy 2016](#_ENREF_27)). In opposition, UCE and AHE advocates have supported the application of the multispecies coalescent ([Edwards et al. 2016](#_ENREF_4)).

In this study, we obtained mixed results from multispecies coalescent analyses. The species tree from dataset 4 had high bootstrap support (Fig. S5) and a high degree of congruence to the ML tree based on concatenation (Fig. 6a-c), whereas the ASTRAL analysis from dataset 2 (Fig. S4) had many nodes that were poorly supported compared to the ML tree based on concatenation (Fig. 4b). The number of loci and the length of each locus was greater for dataset 4 than for dataset 2; both of these factors are known to produce more accurate species tree estimations ([Mirarab et al. 2014](#_ENREF_16)). The ML tree from dataset 2, based on concatenation (Fig. 4b) also did not provide strong support for the backbone of Lepidoptera, but provided strong support for some shallow nodes that were low in the ASTRAL species tree (Figs. 4b, S4). The number and length of loci, as well as the amount of missing data and taxa, are known to make species tree estimation difficult ([Mirarab et al. 2014](#_ENREF_16)), which is especially true for a group thought to have resulted from a rapid radiation.

**Lepidoptera relationships**

Relationships among seven butterfly families in our study were characterized by strong bootstrap support. Our analyses strongly supported the inclusion of Hedylidae and Hesperiidae in the Papilionoidea, a result consistent with recent studies ([Heikkilä et al. 2012](#_ENREF_7); [Kawahara and Breinholt 2014](#_ENREF_10); [Kristensen and Skalski 1998](#_ENREF_13); [Mutanen et al. 2010](#_ENREF_17); [Regier et al. 2013](#_ENREF_22); [Timmermans et al. 2014](#_ENREF_28)). Pieridae, which was not sampled in the study of Kawahara and Breinholt ([2014](#_ENREF_10)), was represented in the present study by *Neophasia terlooii* (433 loci)*.* In nearly all of our analyses, relationships among the seven butterfly families were robust, congruent with Heikkilä et al. ([2012](#_ENREF_7)), and all but two of the 13 nodes had 100% bootstrap support (Figs. 4b, 5a)*.* The ASTRAL tree from dataset 2 failed to resolve relationships among butterfly families (Fig. S4), a result that is likely due to short locus length, the limited number of loci included, and difficulties associated with estimating species trees from gene trees that have missing taxa.

Relationships estimated for Sphingidae from datasets 4-6 are largely congruent with the sanger-based phylogenies of Kawahara et al. ([2009](#_ENREF_11)) and Kawahara and Barber (2015). However, the branching pattern for some groups in prior studies are not consistent with results from the present study. For instance, the Ambulycini, a group for which its placement within Smerinthinae remained equivocal (Kawahara et al. 2009), but is derived within the subfamily based on our new results. Kawahara et al. (2009) recovered Hemarina as the sister-group to all other Macroglossinae, with moderate (~70%) ML bootstrap support. This pattern contradicts the strong support for the placement of *Hemaris* as sister to *Neogurelca* in the ML analysis of datasets 4 and 5 (BS = 98%, Fig. 6a, b) and is also in conflict with the ASTRAL tree analysis of dataset 4 (Fig. S5), where *Neogurelca* is the sister group to other Macroglossinae. The derived position of Hemarina in the Macroglossinae is also strongly supported (BS = 100%) in the study of Breinholt and Kawahara ([2013](#_ENREF_3)). Morphologically, *Hemaris*, *Neogurelca*, and *Sphingonaepiopsis* (the latter which was the sister group to *Neogruelca* in Kawahara et al. ([2011](#_ENREF_8)) and Kawahara and Barber (2015), share crepuscular/diurnal adult diel activity times and small body size (e.g, [Mell 1992](#_ENREF_15); [Rothschild and Jordan 1903](#_ENREF_23); [Tuttle 2007](#_ENREF_29)) and we therefore have stronger confidence for results from the present study. Monophyly and phylogenetic relationships of Saturniidae, inferred in the ML analyses of datasets 4-6 (Fig. 6a-c), are highly congruent with the phylogenetic results of Regier et al. ([2008](#_ENREF_21)), Zwick et al. ([2011](#_ENREF_30)) and Barber et al. ([2015](#_ENREF_1)). The Oxyteninae is the most distant subfamily within the family, and the Ceratocampinae + Hemileucinae are separated from the Saturniini + Attacini (Fig. 6a-c).

**Data set comparison**

The 2,696-locus transcriptome nucleotide alignment of Kawahara and Breinholt ([2014](#_ENREF_10)) had ~28 times more nucleotides, ~48 times more alignment patterns, and ~50 times more informative sites than dataset 2 (557 loci). Even though the 465-locus dataset from Kawahara and Breinholt ([2014](#_ENREF_10)) did not provide strong branch support for the backbone of Lepidoptera, that dataset had approximately five times the number of nucleotides, seven times more alignment patterns, and eight times more informative sites than dataset 2. The percent of sites that were informative (# of informative site/total # of sites) for the 2,696-locus transcriptomic dataset of Kawahara and Breinholt ([2014](#_ENREF_10)), dataset 2 were 37% and 20%, respectively. Datasets 4-6, constructed to examine the efficacy of the Lep1 set for relationships of Bombycoidea, included different regions of the capture sequence. Dataset 4 (probe + flanks) had 35% informative sites, whereas dataset 5 (probe region) and dataset 6 (flanking regions) had 34% and 35% informative sites, respectively. The flanking data does not appear to include significantly more informative sites than the target probe region. The distribution and number of reference taxa included in the Lep1 kit likely played a role in the differences in percent of informative sites between dataset 2 and dataset 5. Including more references throughout the phylogeny would likely allow more efficient capture of loci with higher variation across all of Lepidoptera and could allow more successful enrichment of loci across Lepidoptera with higher amounts of variation.

**References**

Barber JR, Leavell BC, Keener AL, Breinholt JW, Chadwell BA, McClure CJW, Hill GM, Kawahara AY. 2015. Moth tails divert bat attack: Evolution of acoustic deflection. Proceedings of the National Academy of Sciences, 112:2812-2816.

Brandley MC, Bragg JG, Singhal S, Chapple DG, Jennings CK, Lemmon AR, Lemmon EM, Thompson MB, Moritz C. 2015. Evaluating the performance of anchored hybrid enrichment at the tips of the tree of life: a phylogenetic analysis of Australian Eugongylus group scincid lizards. BMC Evol. Biol., 15:1-14.

Breinholt JW, Kawahara AY. 2013. Phylotranscriptomics: Saturated third codon positions radically Influence the estimation of trees based on next-gen data. Genome Biol Evol, 5:2082-2092.

Edwards SV, Xi Z, Janke A, Faircloth BC, McCormack JE, Glenn TC, Zhong B, Wu S, Lemmon EM, Lemmon AR, Leaché AD, Liu L, Davis CC. 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol., 94, Part A:447-462.

Eytan RI, Evans BR, Dornburg A, Lemmon AR, Lemmon EM, Wainwright PC, Near TJ. 2015. Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. BMC Evol. Biol., 15:1-20.

Gatesy J, Springer MS. 2014. Phylogenetic analysis at deep timescales: Unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol., 80:231-266.

Heikkilä M, Kaila L, Mutanen M, Pena C, Wahlberg N. 2012. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. Proc Biol Sci, 279:1093-1099.

Kawahara A, Ohshima I, Kawakita A, Regier J, Mitter C, Cummings M, Davis D, Wagner D, De Prins J, Lopez-Vaamonde C. 2011. Increased gene sampling strengthens support for higher-level groups within leaf-mining moths and relatives (Lepidoptera: Gracillariidae). BMC Evol. Biol., 11:1-14.

Kawahara AY, Barber JR. 2015. Tempo and mode of antibat ultrasound production and sonar jamming in the diverse hawkmoth radiation. Proceedings of the National Academy of Sciences, 112:6407-6412.

Kawahara AY, Breinholt JW. 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. Proceedings of the Royal Society B: Biological Sciences, 281.

Kawahara AY, Mignault AA, Regier JC, Kitching IJ, Mitter C. 2009. Phylogeny and Biogeography of Hawkmoths (Lepidoptera: Sphingidae): Evidence from Five Nuclear Genes. PLoS ONE, 4:e5719.

Kimball RT, Wang N, Heimer-McGinn V, Ferguson C, Braun EL. 2013. Identifying localized biases in large datasets: A case study using the avian tree of life. Mol. Phylogenet. Evol., 69:1021-1032.

Kristensen NP, Skalski AW. 1998. Phylogeny and palaeontology. In: Kristensen NP editor. Handbuch der Zoologie, a Natural History of the Phyla of the Animal Kingdom, Vol. IV, Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies, Vol. 1: Evolution, Systematics, and Biogeography. Berlin & New York, Walter de Gruyter, p. 7-25.

McCormack JE, Harvey MG, Faircloth BC, Crawford NG, Glenn TC, Brumfield RT. 2013. A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing. PLoS ONE, 8:e54848.

Mell R. 1992. Biologie und Systematik der südchinesischen Sphingiden. Berlin, R. Friedländer & Sohn.

Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics, 30:i541-i548.

Mutanen M, Wahlberg N, Kaila L. 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proceedings of the Royal Society B: Biological Sciences.

Peloso PLV, Frost DR, Richards SJ, Rodrigues MT, Donnellan S, Matsui M, Raxworthy CJ, Biju SD, Lemmon EM, Lemmon AR, Wheeler WC. 2016. The impact of anchored phylogenomics and taxon sampling on phylogenetic inference in narrow-mouthed frogs (Anura, Microhylidae). Cladistics, 32:113-140.

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature, 526:569-573.

Pyron RA, Hendry CR, Chou VM, Lemmon EM, Lemmon AR, Burbrink FT. 2014. Effectiveness of phylogenomic data and coalescent species-tree methods for resolving difficult nodes in the phylogeny of advanced snakes (Serpentes: Caenophidia). Mol. Phylogenet. Evol., 81:221-231.

Regier JC, Grant MC, Mitter C, Cook CP, Peigler RS, Rougerie R. 2008. Phylogenetic relationships of wild silkmoths (Lepidoptera: Saturniidae) inferred from four protein-coding nuclear genes. Syst. Entomol., 33:219-228.

Regier JC, Mitter C, Zwick A, Bazinet AL, Cummings MP, Kawahara AY, Sohn J-C, Zwickl DJ, Cho S, Davis DR, Baixeras J, Brown J, Parr C, Weller S, Lees DC, Mitter KT. 2013. A large-scale, higher-level, molecular phylogenetic study of the Insect order Lepidoptera (moths and butterflies). PLoS ONE, 8:e58568.

Rothschild LW, Jordan K. 1903. A revision of the lepidopterous family Sphingidae. Novitates Zoologicae, 9:1-972.

Ruane S, Raxworthy CJ, Lemmon AR, Lemmon EM, Burbrink FT. 2015. Comparing species tree estimation with large anchored phylogenomic and small Sanger-sequenced molecular datasets: an empirical study on Malagasy pseudoxyrhophiine snakes. BMC Evol. Biol., 15:1-14.

Simmons MP, Gatesy J. 2015. Coalescence vs. concatenation: Sophisticated analyses vs. first principles applied to rooting the angiosperms. Mol. Phylogenet. Evol., 91:98-122.

Song S, Liu L, Edwards SV, Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences, 109:14942-14947.

Springer MS, Gatesy J. 2016. The gene tree delusion. Mol. Phylogenet. Evol., 94, Part A:1-33.

Timmermans MJTN, Lees DC, Simonsen TJ. 2014. Towards a mitogenomic phylogeny of Lepidoptera. Mol. Phylogenet. Evol., 79:169-178.

Tuttle J. 2007. The hawk moths of North America: A natural history study of the Sphingidae of the United States and Canada. Washington, D.C., The Wedge Foundation.

Zwick A, Regier JC, Mitter C, Cummings MP. 2011. Increased gene sampling yields robust support for higher-level clades within Bombycoidea (Lepidoptera). Syst. Entomol., 36:31-43.