Supporting Information

Wiegmann et al. 10.1073/pnas.1012675108

SI Materials and Methods

Taxa Sampled, DNA Extraction, Amplification, and Sequencing. Data collection was completed in two tiers. Tier 1, comprising 42 taxa representing the order Diptera and three holometabolous outgroups, was sampled for sequence data from 12 nuclear protein-coding genes, 18S and 28S ribosomal DNA, and complete mitochondrial genomes, along with 371 morphological characters (Table S1). In a few cases, different species were used for mitochondrial sequencing vs. nuclear genomic sequences, but all were sampled from the same family or terminal taxonomic lineage and concatenated as "chimeric" taxa in combined datasets (Table S1). Tier 2, comprising 202 taxa and three additional outgroups, including at least one species from 149 of the ~157 recognized families, was sampled for 5 nuclear genes (7 kb; Table S1). The majority of sequencing work was completed at North Carolina State University, according to the following protocol, but large portions were also performed in the laboratories of J.S. and R.M. following similar methods.

Nuclear genomic DNA was extracted using the DNeasy DNA extraction kit (QIAGEN, Inc.) and the RNA extraction kit (for RT-PCR) (QIAGEN, Inc.). The standard protocol was altered by extending the amount of time the specimen was in proteinase K solution to 2 d to allow enzymes to penetrate the cuticle without grinding the specimen. The final elution was reduced to 30 μL to avoid diluting the DNA solution. Genes were amplified and sequenced using published Diptera primers for 28S and 18S (1) as well as degenerate primers designed by J. K. Moulton for CAD (2) and by J.-W.K. for the remaining protein-coding genes. Primer sequences varied and were often family- or taxon-specific (available on request from B.M.W.). PCR parameters varied for all genes but followed typical three-step reaction protocols (available on request from B.M.W.). PCR products were extracted from agarose gels and purified with the Qiaquick Gel Extraction kit (QIAGEN, Inc.). Big Dye Sequencing kits (Applied Biosystems) were used for sequencing reactions, and sequencing was completed at the North Carolina State University Genomic Sciences Laboratory. Sequences were assembled and edited using Sequencher 4.1 (Gene Codes Corp.). Alignment of ribosomal genes for tier 1 was inferred by eye through manual alignment in the on-screen alignment editor of GDE 2.2 (http://www-bimas.cit.nih.gov/gde_sw.html) and adjusted according to published secondary structure models (3, 4). Alignment of protein-coding genes was carried out manually according to the amino acid translation using Sequence Alignment Editor, version 2.0 (Se-Al 2.0). Introns and other positions of ambiguous alignment were removed from the analysis. The 28S data for tier 2 were aligned in a hierarchical fashion using the program MUSCLE (5) in three separate runs. We first aligned all dipteran sequences (using only tier 1 exemplars for Schizophora), brachyceran sequences only next, and then those of Schizophora only. For each alignment, regions of ambiguous alignment were delimited and excluded using Gblocks (Gblocks version 0.91b) (6), with exclusion regions adjusted manually. The alignments were then combined manually in Se-Al 2.0, such that regions found to be conservative at lower levels (Schizophora only or all Brachycera) but unalignable at higher levels (lower Diptera and/or lower Brachycera and lower Cyclorrhapha) were included for conserved taxon blocks only. To detect existing base compositional bias, a χ^2 test of homogeneity of base frequencies across taxa was performed for the concatenated dataset using Tree Puzzle (7) (and PAUP). For the combined tiers dataset, all three partitions are significantly heterogenous according to PAUP, with or without third positions (ribosomal, mitochondrial, and nuclear), although the mitochondrial partition approaches nonsignificance without third positions ($P = 0.04$).

Mitochondrial Genomes. Mitochondrial genomes were obtained by the laboratories of A.T.B. and M.F. Approximately 1–3 g of tissue was ground in the presence of proteinase K, and total genomic DNA was extracted using a standard phenol–chloroform extraction protocol and Nucleospin DNA purification columns (Macherey–Nagel) later. After ethanol precipitation, extracts were dried and dissolved in 100–200 μL of distilled water. The general strategy for amplification and sequencing was to amplify fragments of 500–1,500 bp using standard primers (8). Details of the amplification conditions and purification of templates are given by Beckenbach and Stewart (9). In the laboratory of M.F., PCR fragments were cleaned up with the QIAquick PCR Purification kit (QIAGEN, Inc.) and sequencing was done with BigDye (Applied Biosystems) and outsourced to MCLab. Purified PCR products were sequenced from both strands using the amplification primers. Additional, taxon-specific primers were designed as needed to fill gaps that occur when standard primers failed to produce a usable product. In the laboratory of M.F., base calling was performed with Phred (Phred; Codoncode Corp.) and contig assembly was done via Phrap/Cross_match/Swat (Phrap; Codoncode Corp.). Contigs were visualized and assembled in BioLign version 4.0.6.2.

Morphology. The analysis of morphological characters was completed by the international team of C.L., B.S., T.P., G.C., J.S., R.M., D.K.Y., A. Borkent, and V.B. A primary obstacle in coding characters across the entire order Diptera was the issue of homology, especially when comparing Cyclorrhapha with lower Diptera (e.g., larval mouthparts, male genitalia). We were initially guided in determination of homology by two major historical works that proposed dipteran relationships based on studies of morphological character systems, Hennig (10) and the *Manual of Nearctic Diptera* (11–13). Hennig (10) discussed the monophyly of major groupings of Diptera. His systematic approach was the foundation for the studies of relationships in the third volume of the *Manual of Nearctic Diptera* (11–13), commencing with relationships of Diptera with other holometabolous orders and within the lower Diptera (11–13) and proceeding with the suborder Brachycera and groupings of the lower Brachycera, and, finally, the Cyclorrhapha. We also considered studies of specific anatomical structures, including antennae (14, 15), male genitalia (16–18), larval mouthparts (19), and wing venation, among orders (20). We added morphological studies using quantitative matrix-based phylogenetic methods examining relationships among the lower or nematocerous families (1–23), lower Brachycera (24), Empidoidea (25, 26), acalyptrates (27), and calyptrates (28, 29). From these numerous studies, we established a list of 457 possible morphological characters to cover the anatomical diversity of Diptera. Over 2 y, team consultation reduced this to a comprehensive list of 371 external and internal morphological characters for larvae (93), pupae (11), and adults (267, including 55 head, 54 wing, 31 female genitalia, and 49 male genitalia). Morphological characters cover the diversity of all fly families and the five holometabolous outgroups, and the list is available on the FLYTREE morphology Web site. We scored 42 first-tier dipteran exemplar taxa and 5 holometabolous outgroup taxa for these characters to produce a morphological supermatrix to examine relationships of Diptera (*a*) with other holometabolous orders, (*b*) major infraordinal level groupings,

(*c*) superfamilies at phylogenetic suture zones, and (*d*) among families. Because only 47 taxa have been scored for the tier 1 analyses, there are some constant characters in the morphological supermatrix.

Phylogenetic Analyses. ***ML analyses.*** Partitioned ML analyses were completed with the molecular data from tier 1, tier 2, and tier 1 and tier 2 combined, all with third positions of protein-coding genes removed (nuclear and mitochondrial), using RAxML 7.0.4 (30) (Fig. S1 and Table S2). We used the GTRMIX option, which optimizes the topology using the GTRCAT approximation and then computes a final likelihood score for comparison using a standard GTR + Γ model. Invariant sites were not modeled, following advice in the RAxML manual. Separate partitions were created for nuclear ribosomal, mitochondrial (not included in tier 2), and nuclear protein-coding genes. For the combined dataset, 1,000 independent runs from random starting trees were performed to find the highest scoring replicate (31). Other analyses were performed with 100 replicates. Node support was calculated by acquiring bootstrap values from heuristic searches of 1,000 resampled datasets, using the rapid bootstrap feature of RAxML (32). Additional exploratory analyses were performed on subsets of the combined dataset, including removal of unstable taxa (see below), removal of Streblidae, and removal of outgroups. The unstable placement of *Nymphomyia* was investigated by selectively pruning taxa possibly involved in long-branch attraction artifacts: all outgroups first and then outgroups plus *Axymyia*. This resulted in the grouping of *Nymphomyia* and *Axymyia* when outgroups were removed (plausibly attributable to long-branch attraction) and the return of *Nymphomyia* to the original position as sister to remaining Diptera, excluding Deuterophlebiidae, when *Axymyia* and outgroups were all removed.

Bayesian analyses. Tier 1. Bayesian analyses of tier 1 data were completed with all positions included, third positions excluded, and combined with morphology with third positions removed (Fig. S2 and Table S2). An appropriate model of nucleotide evolution, GTR + I + Γ in this case, was chosen using Mr. Modeltest (Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden). Using MrBayes (33, 34), analyses were conducted for 20 million generations and trees were sampled every 1,000 generations, with the first 25% discarded as burn-in. Each nuclear gene was treated as a separate partition; however, when third positions were included, each codon position of each gene was treated as a separate partition. Mitochondrial data were not partitioned by gene but were included in their own partition. When third positions were included, each codon position of the mitochondrial data was treated as a separate partition (first, second, and third).

Exploratory Bayesian analysis of the combined tier 1/tier 2 molecular dataset with third positions removed was also performed (GTR + I + Γ), using more Markov chain Monte Carlo chains than specified by default (two runs of 12 chains each) to facilitate convergence. Three partitions were used: nuclear ribosomal, mitochondrial, and nuclear protein-coding. The final convergence statistic for the two runs (SD of split frequencies) was 0.082 when aborted at ∼40 million generations (2 mo). The first 20 million generations were discarded as burn-in.

Stability Analysis. "Rogue" taxa of unstable placement may often obscure support for otherwise robust clades by lowering clade recovery values. To identify unstable taxa and gauge their effect on support values, we tested stability of the placement of each individual taxon in the set of 1,000 ML trees (including 999 suboptimal trees) from the combined data RAxML analysis. We wrote a script in R, using the ape package (35) to prune each single tip sequentially from the set of ML trees and then calculate a stability statistic, *s*, which summarizes the change in partition frequencies across the tree, s, as follows:

$$s = \sum_{j=1}^{N-1} \frac{(x_j - x_j')d_j}{N}$$

where $N-1$ is the number of internal nodes on the pruned phylogeny, x_j and x_j' are partition frequencies for corresponding individual nodes in the original and pruned tree sets, respectively, and d_j is the depth of each node, measured as the maximum number of branches between the node and any of its descendant tips. This measure gives more weight to instability affecting deeper nodes or multiple nodes simultaneously. Bootstrap support values were then recalculated by pruning taxa (n = 19) above an arbitrary threshold stability value ($s > 0.3$) from the original set of bootstrap trees. Our analysis does not consider instability caused by unstable pairs of taxa or small clades; other more computationally intensive methods accounting for this kind of instability (36) failed for our tree set with available computing power (Fig. S1).

Networks. To visualize conflicting phylogenetic signal in our dataset and in our sample of bootstrap trees, we generated a Neighbor Net and a bootstrap consensus showing all conflicting splits in the program SplitsTree (37).

miRNA Analyses. We constructed miRNA libraries for two fly species from tier 1 dataset, *Episyrphus balteatus* (Syrphidae) and *Megaselia scalaris* (Phoridae). Libraries were constructed as described (38). The small RNA libraries were sequenced using the Roche GS-FLX pyrosequencer of the North Carolina State University Genome Sciences Laboratory, and the resulting sequences analyzed with miRMiner (38). In all, we identified 137,059 miRNA sequences in our sample that were assignable to 1 of 159 known miRNA families.

Divergence Time Estimation. Divergence time estimates were calculated for the ingroup-only combined dataset using the penalized likelihood method (39) in r8s, version 1.71. To calibrate divergence time estimates, the split between the Diptera stem group and the nearest outgroup (*Microchorista*) was fixed at 256 Ma, corresponding to the mean age found by Wiegmann et al. (40). The resulting crown group age for extant Diptera (∼246 Ma) corresponds to the earliest known fossil fly, *Grauvogelia* (∼240 Ma) (41), which may represent a crown-group dipteran (42–44). Other minimum age constraints were applied as listed in Table S3, and the clade Schizophora was constrained to a maximum age of 70 Ma, slightly older than the early Tertiary age of the earliest confirmed schizophoran fossils (42). Cross-validation analyses for the optimal smoothing parameter failed, so final penalized likelihood runs were completed using values of s = 1 and s = 1,000 to obtain a range of plausible divergence times (Fig. S3 and Table S3).

This analysis was repeated without the maximum constraint for Schizophora to gauge the dependency of results on this constraint. This resulted in an inferred age for Schizophora of ∼115 Ma. Discrepancies between molecular- and fossil-based age estimates for major radiations are common and probably reflect frequent violation of molecular-evolutionary assumptions (e.g., correlation of rates on ancestor and descendant branches) as well as a missing fossil record of early stages of radiations. We consider the results with the Schizophora constraint to be more realistic and a missing fossil record for Schizophora of 50 Ma to be unlikely, considering the dipteran diversity present in known Cretaceous deposits (43, 44).

Tests of Diversification. A family-level chronogram for flies was produced by pruning the r8s (s = 1) phylogeny such that each family is represented by a single tip, including three families (Conopidae, Rhagionidae, and Heleomyzidae) that did not ap-

pear as monophyletic. In addition, the representative of Streblidae was pruned, because placement of this family outside of Hippoboscoidea contradicts earlier robust results (45) and may be attributable to systematic error. Approximate diversities of each family were obtained from the current working version of the Systema Dipterorum (46). Diversities for Anthomyzidae and Heleomyzidae were combined, because the latter was found to be paraphyletic with respect to Anthomyzidae. Diversities for Streblidae and Nycteribiidae were likewise added to those of the closely related Hippoboscidae. These data were used as input for the MEDUSA program (47), which sequentially adds breakpoints to a multirate birth–death model fitting the given branch lengths and terminal diversities until subsequent breakpoints do not result in a threshold difference in Akaike Information Criterion (AIC) values (here, $\Delta AIC = 2$). This method estimated a highly significant ($\Delta AIC = 193$) shift in rates at the origin of Schizophora, as well as eight other shifts with much lower but significant ΔAIC values. Estimated breakpoints were similar (7 of 10 breakpoints were identical) for chronograms constructed with different smoothing parameters (Table S4); breakpoints estimated from the chronogram with no maximum constraint for Schizophora were also very similar (8 of 9 breakpoints were identical) to the constrained results, although estimated speciation rates were lower.

To visualize the relationship between clade age and clade size for dipteran families, these were plotted on a semilog scale and compared with expected values under constant rates of speciation and extinction computed using the Geiger packages (48), according to Magallón and Sanderson (49). This was repeated for three values of relative extinction rate (0, 0.9, and 0.99), with overall diversification rates estimated from the data. Under a constant birth–death model, a number of families were either more or less diverse than expected. Regression analysis revealed a lack of correlation between age and log clade size ($P = 0.35$, $r^2 = 0.006$), implying that differences in family diversities may be attributable more to extrinsic limits ("carrying capacities") on clade diversity following initial species radiations than to intrinsic consistent differences in diversification rate (50) (Fig. S4). A similar lack of correlation was found within two of the major rate divisions identified in the MEDUSA analysis ($P = 0.429$, $r^2 = 0.042$ for lower Diptera; $P = 0.592$, $r^2 = 0.016$ for non-eremoneuran Orthorrhapha), but a marginal correlation was found for the other two major rate divisions ($P = 0.057, r^2 = 0.149$ for Neodiptera, excluding Orthorrhapha and Schizophora; $P = 0.048$ and $r^2 = 0.068$ for Schizophora, excluding those lineages with subsequent significant shifts). Unexpectedly, the significant correlation in one case was negative (Neodiptera), implying that older clades are, on average, smaller than younger ones.

1. Yang L, Wiegmann BM, Yeates DK, Irwin ME (2000) Higher-level phylogeny of the Therevidae (Diptera: insecta) based on 28S ribosomal and elongation factor-1 alpha gene sequences. *Mol Phylogenet Evol* 15:440–451.
2. Moulton JK, Wiegmann BM (2004) Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged Eremoneuran Diptera (Insecta). *Mol Phylogenet Evol* 31:363–378.
3. Kjer KM (2004) Aligned 18S and insect phylogeny. *Syst Biol* 53:506–514.
4. Bertone MA, Courtney GW, Wiegmann BM (2008) Phylogenetics and temporal diversification of the earliest true flies (Insecta: Diptera) based on multiple nuclear genes. *Syst Entomol* 33:668–687.
5. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
6. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552.
7. Strimmer K, von Haeseler A (1997) Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci USA* 94:6815–6819.
8. Simon C, Buckley TR, Frait F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu Rev Ecol Evol Syst* 37:545–579.
9. Beckenbach AT, Stewart JB (2009) Insect mitochondrial genomics 3: The complete mitochondrial genome sequences of representatives from two neuropteroid orders: A dobsonfly (order Megaloptera) and a giant lacewing and an owlfly (order Neuroptera). *Genome* 52:31–38.
10. Hennig W (1973) Insecta. 2 Teil: Spezielles. 31. *Handbuch der Zoologie*, eds Helmcke JG, Starck D, Wermuth H (De Gruyter, Berlin) Band IV, Hälfte 2, pp. 337.
11. Wood DM, Borkent A (1989) *Manual of Nearctic Diptera,* Research Branch Canada Agriculture Monograph No. 32, ed McAlpine J (Canadian Government Publishing Centre, Hull, Canada), pp 1333–1370.
12. Woodley N (1989) *Manual of Nearctic Diptera,* Research Branch Agriculture Monograph No. 32, ed McAlpine JF (Canadian Government Publishing Centre, Hull, Canada), Vol 3, pp 1371–1395.
13. McAlpine J (1989) *Manual of Nearctic Diptera,* Research Branch Agriculture Monograph No. 32, ed McAlpine JF (Canadian Government Publishing Centre, Hull, Canada), pp 1397–1581.
14. Stuckenberg BR (1999) Antennal evolution in the Brachycera (Diptera), with reassessment of terminology relating to the flagellum. *Stud Dipterologica* 6:33–48.
15. Stuckenberg BR (2001) Pruning the tree: A critical review of classifications of the Homeodactyla (Diptera, Brachycera), with new perspectives and an alternative classification. *Stud Dipterologica* 8:3–42.
16. Wood DM (1991) Homology and phylogenetic implications of male genitalia in Diptera. The ground plan. *Proceedings of the Second International Congress of Dipterology*, eds Weismann L, Országh I, Pont AC (SPC Academic Publishing, The Hague), pp 255–284.
17. Sinclair BJ, Cumming JM, Wood DM (1994) Homology and phylogenetic implications of male genitalia in Diptera—Lower Brachycera. *Entomol Scand* 24:407–432.
18. Cumming JM, Sinclair BJ, Wood DM (1995) Homology and phylogenetic implications of male genitalia in Diptera—Eremoneura. *Entomol Scand* 26:121–151.
19. Courtney GW, Sinclair BJ, Meier R (2000) *Contributions to a Manual of Palaearctic Diptera (with special reference to flies of economic importance). General and Applied Dipterology*, eds Papp L, Darvas B (Science Herald, Budapest), Vol 1, pp 85–163.
20. Wootton RJ, Ennos AR (1989) The implications of function on the origin and homologies of the dipterous wing. *Syst Entomol* 14:507–520.
21. Oosterbroek P, Courtney G (1995) Phylogeny of the nematocerous families of Diptera (Insecta). *Zool J Linn Soc* 115:267–311.
22. Saether O (2000) Phylogeny of the Culicomorpha (Diptera). *Syst Entomol* 25:223–234.
23. Sinclair BJ, Borkent A, Wood DM (2007) The male genital tract and aedeagal components of the Diptera with a discussion of their phylogenetic significance. *Zool J Linn Soc* 150:711–742.
24. Yeates DK (2002) Relationships of extant lower Brachycera (Diptera): A quantitative synthesis of morphological characters. *Zool Scr* 31:105–121.
25. Wiegmann BM, Mitter C, Thompson FC (1993) Evolutionary origin of the Cyclorrhapha: Tests of alternative morphological hypotheses. *Cladistics* 9:41–81.
26. Sinclair BJ, Cumming JM (2006) The morphology, higher-level phylogeny and classification of the Empidoidea (Diptera). *Zootaxa* 1180:1–172.
27. Buck M (2006) A new family and genus of acalyptrate flies from the neotropical region, with a phylogenetic analysis of Carnoidea family relationships (Diptera, Schizophora). *Syst Entomol* 31:377–404.
28. Pape T (1992) Phylogeny of the Tachinidae family-group (Diptera: Calyptratae). *Tijdschr Entomol* 135:43–86.
29. Rognes K (1997) The Calliphoridae (blowflies) (Diptera: Oestroidea) are not a monophyletic group. *Cladistics* 13:27–66.
30. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
31. Regier JC, et al. (2008) Resolving arthropod phylogeny: Exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequence. *Syst Biol* 57:920–938.
32. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57:758–771.
33. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
34. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
35. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
36. Thorley JL, Page RD (2000) RadCon: Phylogenetic tree comparison and consensus. *Bioinformatics* 16:486–487.
37. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254–267.
38. Wheeler BM, et al. (2009) The deep evolution of metazoan microRNAs. *Evol Dev* 11: 50–68.
39. Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol Biol Evol* 19:101–109.
40. Wiegmann BM, et al. (2009) Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol* 7:34.
41. Krzeminski W, Krzeminska E, Papier F (1994) *Grauvogelia arzvilleriana* sp. n.—The oldest Diptera species (Lower/Middle Triassic of France). *Acta Zool Cracov* 37: 95–99.
42. Winkler IS, Labandeira C, Wappler T, Wilf P (2010) Distinguishing Agromyzidae (Diptera: Schizophora) leaf mines in the fossil record: New taxa from the Paleogene of North America and Germany and their evolutionary implications. *J Paleontol* 84: 935–954.
43. Blagoderov VA, Lukashevich ED, Mostovski MB (2002) *History of Insects*, eds Rasnitsyn AP, Quicke DLJ (Kluwer Academic, Dordrecht), pp 227–240.
44. Grimaldi D, Engel MS (2005) *Evolution of the Insects* (Cambridge Univ Press, New York).
45. Petersen FT, Meier R, Kutty SN, Wiegmann BM (2007) The phylogeny and evolution of host choice in the Hippoboscoidea (Diptera) as reconstructed using four molecular markers. *Mol Phylogenet Evol* 45:111–122.

46. Evenhuis NL, Pape T, Pont AC, Thompson FC eds (2008) Biosystematic database of world Diptera, working version. Available at http://www.diptera.org. Accessed December 5, 2008.
47. Alfaro ME, et al. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci USA* 106:13410–13414.
48. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24:129–131.
49. Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.
50. Rabosky DL (2009) Ecological limits on clade diversification in higher taxa. *Am Nat* 173:662–674.

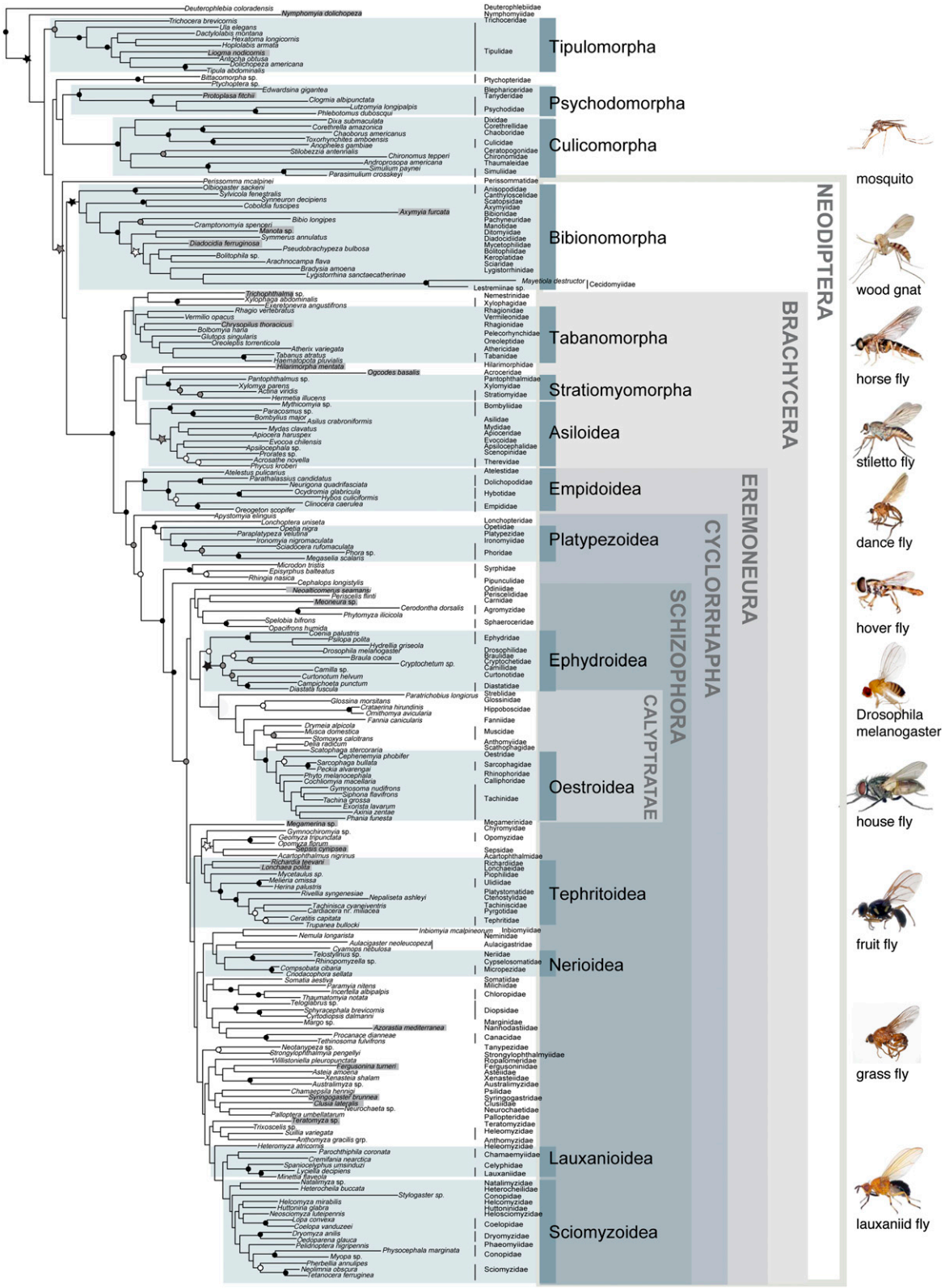

Fig. S1. Tier 1/tier 2 molecular topology, all taxa. Phylogram of ML analysis of the combined tier 1/tier 2 molecular datasets, including all taxa. Circles indicate bootstrap support >80% (black/bp = 95–100%, gray/bp = 88–94%, white/bp = 80–88%). Unstable taxa are gray. Nodes with improved bootstrap values resulting from postanalysis pruning of unstable taxa are marked by stars (black/bp = 95–100%, gray/bp = 88–94%, white/bp = 80–88%).

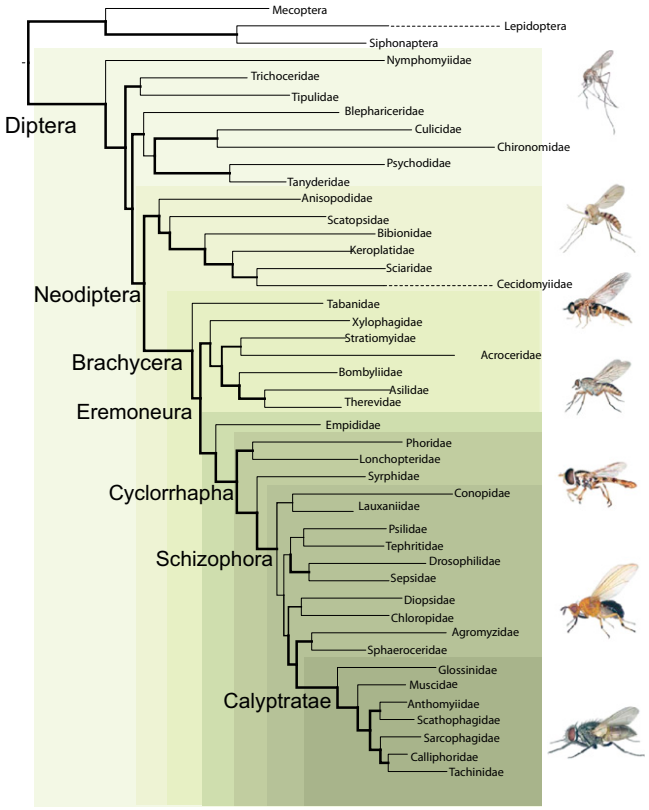

Fig. S2. Tier 1 topology, molecular plus morphology. Phylogram of the Bayesian likelihood analysis of combined molecular and morphological data for the tier 1 taxon set, based on concatenation of 14 nuclear genes, full mitochondrial genomes, and 371 morphological features. Thickened branches indicate nodes supported by >90% posterior probability in Bayesian analysis. Dashed branches are abbreviated to fit within figure dimensions. Outgroup rooting below Diptera is arbitrary.

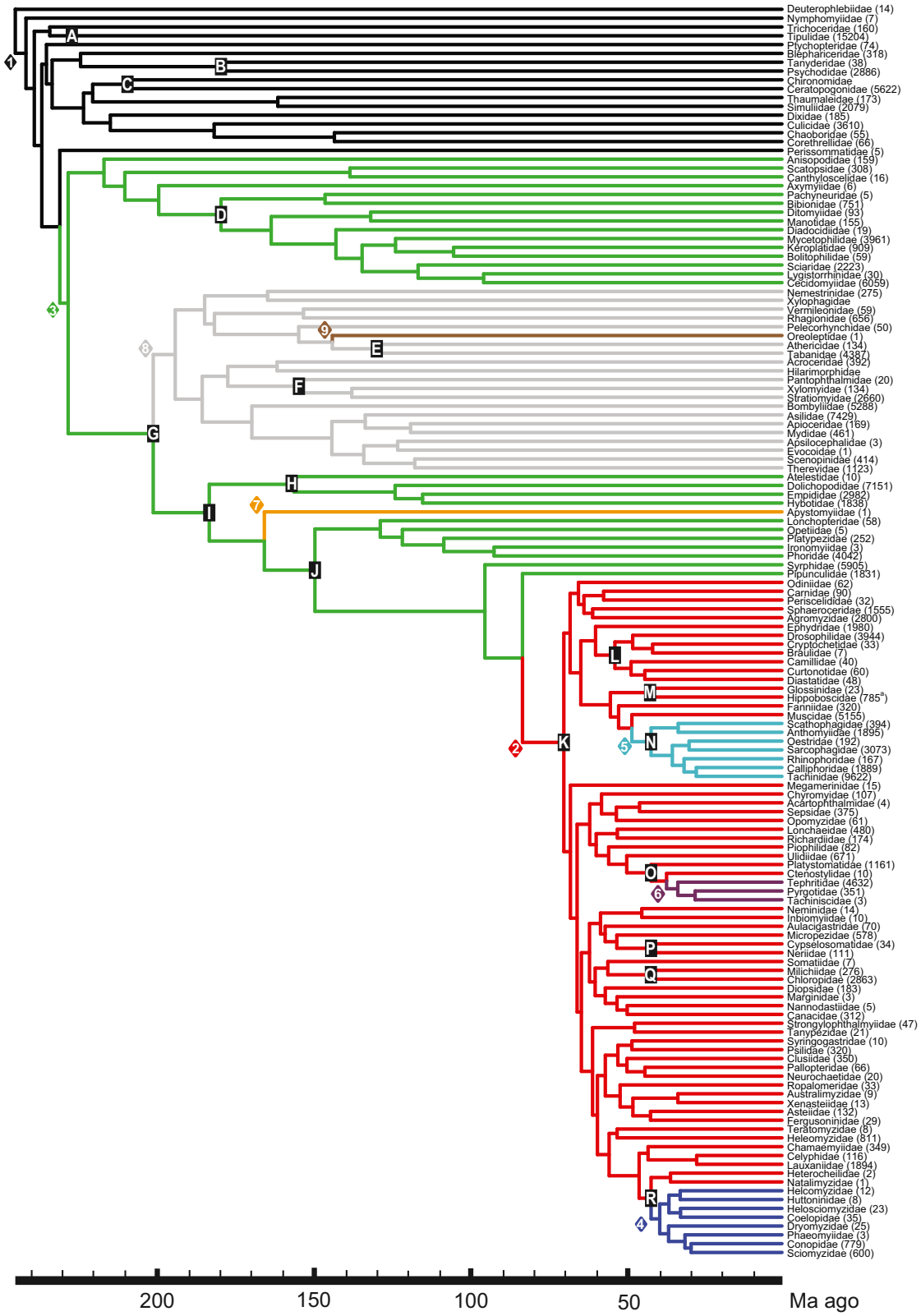

Fig. S3. Complete family-level chronogram. Full penalized likelihood chronogram for dipteran families (smoothing parameter, s = 1.0). Approximate number of described species for each family is listed in parentheses to the right. Fossil constraints applied in the penalized likelihood analysis are marked by letters in black bars (Table S3). Significant shifts in speciation/extinction rates identified in MEDUSA (for s = 1.0 chronogram) are marked by numbered diamonds. Taxa were pruned from the actual results such that each family (even the three that were not monophyletic) is represented by a single taxon; the exemplar for Streblidae, placement of which is possibly artifactual, was also pruned. Thus, the totals for Hippoboscidae include the related Streblidae and Nycteribiidae, and the totals for Heleomyzidae include Anthomyzidae.

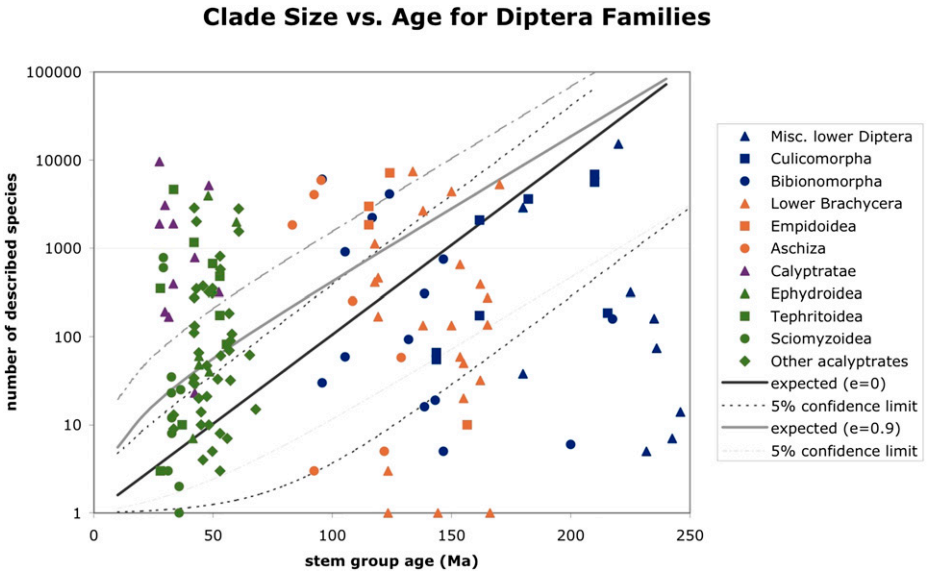

Fig. S4. Plot of species diversity vs. age for extant dipteran families.

Other Supporting Information Files

Table S1 (DOC)
Table S2 (DOC)
Table S3 (DOC)
Table S4 (DOC)