Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 701–713
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

The phylogenetic relationships of flies in the superfamily Empidoidea (Insecta: Diptera)

John K. Moulton [a,*], Brian M. Wiegmann [b]

[a] Department of Entomology and Plant Pathology, 2431 Joe Johnson Drive, 205 Ellington Plant Sciences Building, The University of Tennessee, Knoxville, TN 37996-4560, USA
[b] Department of Entomology, Gardner Hall, Campus Box 7613, North Carolina State University, Raleigh, NC 27695-7613, USA

Received 9 August 2005; revised 21 February 2007; accepted 27 February 2007
Available online 13 March 2007

Abstract

We conducted a molecular phylogenetic study of the Empidoidea, a diverse group of 10,000 species of true flies, with two major goals: to reconstruct a taxonomically complete and robustly supported phylogeny for the group and to use this information to assess several competing classifications for the clade. We amassed 3900+ nucleotides of coding data from the carbamoylphosphate synthase domain of the rudimentary locus (CAD) and 1200+ nucleotides from the large nuclear ribosomal subunit (28S) from 72 and 71 species, respectively, representing several orthorrhaphan and cyclorrhaphan families and all previously recognized empidoidean subfamilies. Independent and combined phylogenetic analyses of these data were conducted using parsimony, maximum likelihood, and Bayesian criteria. The combined matrix included 61 taxa for which both CAD and 28S sequences were obtained. Analyses of CAD first and second codon positions alone and when concatenated with 28S sequences yielded trees with similar and largely stable topologies. Analyses of 28S data alone supported many clades although resolution is limited by low sequence divergence. The following major empidoid clades were recovered with convincing support in a majority of analyses: Atelestidae, Empidoidea exclusive of Atelestidae, Hybotidae sensu lato, Dolichopodidae + Microphorinae (including Parathallassius), and Empididae sensu lato (including Brachystomatinae, Ceratomerinae, Clinocerinae, Empidinae, Hemerodromiinae, Oreogetoninae, and Trichopezinae). The branching arrangement among these four major clades was Atelestidae, Hybotidae, Dolichopodidae/Microphorinae, Empididae. Previously recognized subclades recovered with robust support included Hybotinae, Brachystomatinae, Tachydromiinae, Clinocerinae (in part), Hemerodromiinae, Empidinae, and Empidiini.
Published by Elsevier Inc.

Keywords: Diptera; Dance fly; Phylogeny; Systematics; DNA; Protein

1. Introduction

Chvála (1983) proposed that Empidoidea should be divided into four family-level group taxa—Atelestidae, Hybotidae, Microphoridae + Dolichopodidae, and Empididae *s.str.*—based on morphological grounds. This classification has not been universally accepted, however, particularly in regions outside of Europe, with most workers electing to retain the classical two family system, i.e., Dolichopodidae and Empididae. This action creates a heterogeneous Empididae that is comprised of several subfamilies having dubious monophyly (Cumming et al., 1995; Sinclair and Cumming, 2006). Great morphological and species diversity, along with competing alternative interpretations of homology especially in the male genitalia, have perpetuated the differential use of rival classifications (Chvála, 1983; Cumming et al., 1995; Sinclair, 1995; Sinclair and Cumming, 2006).

Six classifications for the Empidoidea based on phylogenetic hypotheses have been published since Chvála (1983)—Wiegmann et al. (1993), Cumming et al. (1995), Collins and Wiegmann (2002a), Moulton and Wiegmann (2004), Yang (2004), and, most recently, Sinclair and

* Corresponding author. Fax: +1 865 974 4744.
E-mail address: jmoulton@utk.edu (J.K. Moulton).

1055-7903/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.ympev.2007.02.029

Cumming (2006). Wiegmann et al. (1993); Cumming et al. (1995), Yang (2004), and Sinclair and Cumming (2006) are based on morphological evidence while the others were inferred from nucleotide sequence data. Only two superfamilial clades—Microphorinae (='idae' of Chvála, 1983) + Dolichopodidae and Hybotidae (=Hybotinae + Ocydromiinae + Tachydromiinae)—are common to all of these classifications and to that of Chvála (1983). The greatest degree of concordance among the competing classifications/phylogenies exists between those of Chvála (1983) and Moulton and Wiegmann (2004). Moulton and Wiegmann (2004) recovered trees largely concordant with Chvála's (1983) empidoidean classification, with the exception being that Chvála inferred Atelestidae to be the sister group to Cyclorrhapha. Both studies support the monophyly of Empidoidea exclusive of Atelestidae, of Microphoridae + Dolichopodidae, and of Hybotidae They also support similar arrangements within a monophyletic Empididae *s.str*., although with several groups unsampled, i.e., Oreogetoninae + (Clinocerinae + Empididae). A recent re-analysis of the morphological evidence by Sinclair and Cumming (2006) proposes five monophyletic families, Atelestidae, Hybotidae, Dolichopodidae (including Microphorinae s.l.), Empididae and Brachystomatidae. The latter newly proposed family is a clade comprising former brachystomatine, ceratomerine, and trichopezine empidids (Sinclair and Cumming, 2006).

Concordance of phylogenetic trees of Eremoneura inferred from CAD (Moulton and Wiegmann, 2004), 28S rDNA (Collins and Wiegmann, 2002a), and morphological data (Chvála, 1983) was used to demonstrate the phylogenetic utility of CAD for reconstructing Mesozoic-aged divergences. We now build upon our previous molecular phylogenetic investigations of Empidoidea by dramatically increasing taxon sampling for both CAD and 28S rDNA and by conducting independent and combined phylogenetic analyses of these data. Our objective was to obtain a robust molecular phylogeny with which to assess several previously proposed empidoid classifications and, ultimately, to provide a well-sampled phylogenetic framework on which interpretations of empidoid morphological and ecological diversification can be further clarified.

2. Materials and methods

2.1. Taxon sampling

2.1.1. CAD

The ingroup, Empidoidea, is comprised of 55 sampled taxa representing all previously recognized major empidoidean clades, including Atelestidae, Brachystomatinae, Ceratomerinae, Clinocerinae, Dolichopodidae, Empidinae, Hemerodromiinae, Hybotinae, Microphorinae, Ocydromiinae, Oreogetoninae, Tachydromiinae, and Trichopezinae (Griffiths, 1972; Chvála, 1983; Cumming et al., 1995; Disney, 1994; Zatwarnicki, 1996; Sinclair, 1995; Collins and Wiegmann, 2002a,b). Seventeen taxa, representing 4 orthorrhaphan and 10 cyclorrhaphous brachyceran families, served as outgroups, with *Heterostomus curvipalpus* (Xylophagidae) selected as a distal orthorrhaphous brachyceran root.

2.1.2. 28S

A total of 75 taxa were sampled. These taxa represent all of the major empidoid lineages and most of the same outgroup families that were sampled for CAD. The therevids *Ruppellia* and *Efflatouniella* served as the distal root in 28S analyses. Table 1 is a list of the taxa sampled in this study and accession numbers for the CAD and 28S sequences derived from them.

2.2. DNA extraction and amplification

Total genomic nucleic acids were extracted from whole specimens or dissected tissues from fresh, frozen, or ethanol-preserved individuals. Voucher material is deposited in the laboratory of BMW. Nucleotide extractions were carried out using a guanidinium isothiocyanate (Chirgwin et al., 1979) or sodium dodecyl sulfate/proteinase K-based homogenization protocol. Homogenates were extracted once with a 25:24:1 mixture of phenol:chloroform:isoamyl alcohol (Sigma–Aldrich) and again with a 24:1 chloroform:isoamyl alcohol mixture. The DNA was salted out using one-tenth volume of 3 M sodium acetate and 1× volume of chilled (−20 °C) isopropanol and then pelleted by microcentrifugation. DNA was washed with 70 and 95% ethanol, air-dried, resuspended in 100 μl of 1 mM Tris–EDTA, and stored at −80 °C.

The oligonucleotide primers used in this study to amplify CAD and 28S are shown in Table 2. Detailed information about the design of CAD primers are outlined in Moulton and Wiegmann (2004). The primers used to amplify 28S were modified from ones developed for plants to match the published *Drosophila melanogaster* sequence (Hancock et al., 1988; Collins and Wiegmann, 2002a). CAD fragments were amplified using ExTaq Hot Start polymerase (TakaraMirus Bio, Madison, WI), and most 28S amplifications were facilitated by *Taq* polymerase from Promega Biotech, Maddison, WI. Some 28S amplicons were obtained via reverse transcriptase PCR using the GeneAmp RNA PCR Kit (PE Applied Biosystems, Foster City, CA). Specific details about 28S and CAD amplification parameters appear in Collins and Wiegmann (2002a) and Moulton and Wiegmann (2004). DNA amplifications were qualified via horizontal electrophoresis in agarose. CAD amplicons were excised from agarose gels and purified with QiaQuick Gel Extraction Kits (Qiagen Inc., Santa Clara, CA) using the manufacturer's suggested protocol. 28S amplicons were purified using QiaQuick PCR Purification Kits (Qiagen, Inc., Santa Clara, CA) using the manufacturer's suggested protocol. Purified PCR fragments were quantified via horizontal electrophoresis in agarose via comparison to a band of known concentration in a standard DNA ladder.

Table 1
Taxa included in this study

Taxon	Family	Geographic Origin	CAD GenBank Accession Nos.	28S GenBank Accession Nos.
Orthorrhaphans				
Heterostomus curvipalpus	Xylophagidae	CHILE	AY280682	N/A
Bombylius major	Bombyliidae	USA (NC)	AY280675	N/A
Mythicomyia sp.	Bombyliidae	USA (CA)	AY267613	N/A
Hilarimorpha sp.	Hilarimorphidae	USA (UT)	AY280683	N/A
Agapophytus bicolor	Therevidae	AUSTRALIA (QLD)	AY280673	N/A
Efflatouniella aegyptiaca	Therevidae	ISRAEL	AY280679	DQ496211
Ruppellia basalis	Therevidae	NAMIBIA	N/A	AF503031, AF503076
Eremoneura				
Cyclorrhapha				
Aschiza				
Lonchoptera bifurcata	Lonchopteridae	USA (NC)	AY280687	DQ508417
Phora sp.	Phoridae	USA (NC)	AY280694	DQ496183
Opetia nigra	Opetiidae	EU (U. K.)	AY280692	AF502991, AF503012
Paraplatypeza atra	Platypezidae	EU (U. K.)	AY280693	AF502993, AF503014
Pipunculus houghi	Pipunculidae	USA (NC)	AY280691	DQ496182
Ocyptamus sp.	Syrphidae	USA (NC)	DQ369280	N/A
Rhingia nasica	Syrphidae	USA (PA)	AY280697	AF502998, AF503019
Schizophora				
Ceroxys edwardsi	Ulidiidae	Chile	N/A	AF503002, AF503023
Chaetopsis sp.	Ulidiidae	USA (NC)	AY280676	N/A
Pherbellia cinerella	Sciomyzidae	France	N/A	DQ508418
Drosophila melanogaster	Drosophilidae	Unknown	AAF48639	M21017
Epalpus signifer	Tachinidae	USA (AZ)	AY280680	DQ496193
Musca domestica	Muscidae	USA (AZ/NCSU Lab)	AY280689	AF503004, AF503025
Empidoidea				
Acarteroptera recta	Atelestidae	CHILE	AY280672	AF503032, AF503077
Atelestus pulicarius	Atelestidae	EU (U. K.)	AY280700, AY280701	AF503033, AF503078
Meghyperus occidens	Atelestidae	USA (CA)	DQ369300	AF503035, AF503080
Meghyperus sudeticus	Atelestidae	EU (LUXEMBOURG)	AY280688	AF503034, AF503079
Campsicnemus curvipes	Dolichopodidae	EU (CZECH REP.)	DQ369278	DQ496188
Dolichopus sp. "3"	Dolichopodidae	USA (WA)	AY280678	AF503069, AF503114
Nematoproctus sp.	Dolichopodidae	USA (NC)	DQ369304	DQ496190
Neurigona quadrifasciata	Dolichopodidae	EU (CZECH REP.)	AY280690	DQ496199
Poecilobothris nobilitatus	Dolichopodidae	EU (CZECH REP.)	DQ369296	DQ496206
Brachystoma robertsonii	Empididae: Brachystomatinae	USA (NC)	DQ369270	AF503036, AF 503081
Brachystoma sp. "Chile"	Empididae: Oreogetoninae	CHILE (Chiloe Isld.)	DQ369269	DQ496185
Ceratomerus sp.	Empididae: Ceratomerinae	AUSTRALIA	DQ369277	AF503037, AF503082
Clinocera sp. "CA1"	Empididae: Clinocerinae	USA (CA)	AY280677	AF503038, AF503083
Clinocera sp. "CA2	Empididae: Clinocerinae	USA (CA)	DQ369271	DQ496189
Proagomyia sp.	Empididae: Clinocerinae	AUSTRALIA	DQ369302	AF503039, AF503084
Trichoclinocera sp	Empididae: Clinocerinae	USA (TN)	DQ369282	AF503041, AF503086
Wiedemannia sp.	Empididae: Clinocerinae	USA (CA)	AY280699	DQ508416
Deuteragonista sp.	Empididae: Empidinae	CHILE (Chiloe Isld.)	DQ369285	AF503046, AF503091
Empidadelpha sobrina	Empididae: Empidinae	CHILE (Chiloe Isld.)	DQ369275	DQ496192
Empis (*Enoplempis*) sp. "1"	Empididae: Empidinae	USA (TN)	N/A	AF503042, AF503087

(*continued on next page*)

Table 1 (*continued*)

Taxon	Family	Geographic Origin	CAD GenBank Accession Nos.	28S GenBank Accession Nos.
Empis (*Enoplempis*) sp. "25"	Empididae: Empidinae	USA (NC)	AY280681	AF503043, AF503088
Hilara lugubris	Empididae: Empidinae	EU (U. K.)	DQ369299	AF503044, AF503089
Hilara sp. "52"	Empididae: Empidinae	USA (NC)	N/A	DQ496180, DQ496181
Hilara sp. "1"	Empididae: Empidinae	EU (U. K.)	DQ369290	AF503047, AF503092
Hilarigona sp.	Empididae: Empidinae	CHILE (Chiloe Isld.)	DQ369272	DQ496194
Rhamphomyia sp. "Lake"	Empididae: Empidinae	USA (NC)	DQ369297	N/A
Rhamphomyia sp. "3"	Empididae: Empidinae	USA (NC)	N/A	AF503048, AF503093
Rhamphomyia sp. "38"	Empididae: Empidinae	USA (PA)	AY280696	AF503049, AF503094
Chelifera sp.	Empididae: Hemerodromiinae	USA (NC)	DQ369298	N/A
Chelipoda sp.	Empididae: Hemerodromiinae	AUSTRALIA	DQ369268	AF503050, AF503095
Metachela collusor	Empididae: Hemerodromiinae	USA (UT)	DQ369295	DQ496198
Anthepiscopus sp.	Empididae: Oreogetoninae	USA (UT)	DQ369283	DQ496187
Hesperempis sp.	Empididae: Oreogetoninae	USA (CO)	DQ369279	AF503053, AF503098
Hormopeza sp.	Empididae: Oreogetoninae	USA (MN)	DQ369306	DQ496209, DQ496210
Iteaphila sp.	Empididae: Oreogetoninae	USA (UT)	AY280685	DQ496197
Oreogeton scopifer	Empididae: Oreogetoninae	USA (NV)	DQ369287	DQ496201
Apalocnemis sp. "1"	Empididae: Trichopezinae	AUSTRALIA	DQ369266	AF503051, AF503096
Apalocnemis sp. "2"	Empididae: Trichopezinae	NEW ZEALAND	N/A	AF503052, AF503097
Apalocnemis sp. "Chile"	Empididae: Trichopezinae	CHILE (Chiloe Isld.)	DQ369273	N/A
Gloma sp.	Empididae: Trichopezinae	USA (CA)	EF539207	N/A
Heleodromia sp.	Empididae: Trichopezinae	USA (UT)	DQ369303	DQ496184
Heterophlebus sp. "Plain"	Empididae: Trichopezinae	CHILE (Chiloe Isld.)	DQ369292	DQ496202
Heterophlebus sp. "Stripe"	Empididae: Trichopezinae	CHILE (Chiloe Isld.)	DQ369291	DQ496203
Hyperperacera nemoralis	Empididae: Trichopezinae	CHILE (Chiloe Isld.)	DQ369294	DQ496196
Trichopeza longicornis	Empididae: Trichopezinae	EU (GERMANY)	DQ369289	DQ496207
Euhybus sp.	Hybotidae: Hybotinae	EU (U. K.)	DQ369293	AF503054, AF503099
Hybos sp.	Hybotidae: Hybotinae	AUSTRALIA	AY280684	DQ496195
Syneches sp.	Hybotidae: Hybotinae	USA (NC)	DQ369276	AF503055, AF503100
Anthalia sp.	Hybotidae: Ocydromiinae	USA (CA)	AY280674	AF503056, AF503101
Bicellaria sp.	Hybotidae: Ocydromiinae	USA (NC)	DQ369267	AF503057, AF503102
Euthyneura sp.	Hybotidae: Ocydromiinae	USA (NC)	DQ369274	AF503060, AF503105
Hoplopeza sp.	Hybotidae: Ocydromiinae	AUSTRALIA	DQ369286	AF503058, AF503103
Leptopeza sp.	Hybotidae: Ocydromiinae	USA (NC)	AY280686	AF503059, AF503104
Ocydromia glabricula	Hybotidae: Ocydromiinae	EU (GERMANY)	DQ369301	DQ496200
Oropezella sphenoptera	Hybotidae: Ocydromiinae	EU (GERMANY)	DQ369288	DQ496204
Trichinomyia flavipes	Hybotidae: Ocydromiinae	EU (GERMANY)	DQ369281	DQ496186
Drapetis sp.	Hybotidae: Tachydrominae	USA (NC)	DQ369284	DQ496191
Platypalpus sp.	Hybotidae: Tachydromiinae	USA (NC)	AY280695	AF503063, AF503108
Schistostoma sp.	Microphoridae: Microphorinae	USA (CA)	AY280698	AF503066, AF503110
Amphithalassius latus	Microphoridae: Parathalassinae	S. AFRICA	N/A	DQ508415
Parathalassius sp.	Microphoridae: Parathalassinae	USA (CA)	DQ369305	DQ496205

Table 2
Oligonucleotide primers used in this study to amplify and sequence CAD and 28S

Primer	Position[c]	Length	Sequence (5′ to 3′)
CAD[a,b]			
54F	112	23	GTN GTN TTY CAR ACN GGN ATG GT
364R	950	23	TCN ACN GCR AAN CCR TGR TTY TG
405R	1073	23	GCN GTR TGY TCN GGR TGR AAY TG
320F	793	26	ATH TTY GGN ATY TGY YTG GGN CAY CA
338F	856	26	ATG AAR TAY GGY AAT CGT GGH CAY AA
680R	1856	26	AAN GCR TCN CGN ACM ACY TCR TAY TC
581F2	1459	26	GGW GGW CAA ACW GCW YTM AAY TGY GG
581F	1561	26	GAN ACT GAR GAY MGR AAA ATM TTY GC
843R	2369	23	GCY TTY TGR AAN GCY TCY TCR AA
843R2	2384	26	TCN ACC ATW CKN ARW GCY TTY TGR AA
787F	2200	26	GGD GTN ACN ACN GCN TGY TTY GAR CC
806F	2242	23	GTN GTN AAR ATG CCN MGN TGG GA
1098R	3104	20	TTN GGN AGY TGN CCN CCC AT
1124R	3204	27	CAT NCG NGA RAA YTT RAA RCG ATT YTC
1057F	2977	23	GTN TCN ACN GAY TAY GAY ATG TG
1278R	3668	26	TCR TTN TTY TTW GCR ATY AAY TGC AT
1201F	3433	26	GAR GCN AAR GAR ATY GAY GTN GAY GC
1436R	4109	20	CCR TGY TCN GCR TAR AAR TC
28S			
rc28D	5421	18	CCG CAG CTG GTC TCC AAG
28H	5982	18	GGT TTC GCT GGA TAG TAG
rc28H	5999	18	CTA CTA TCC AGC GAA ACC
28K	6582	18	CTT CGA TGT CGG CTC TTG
28Q	6389	18	AAC TCC CTA CCT GGC AAT
rc28Q	6406	22	GGA CAT TGC CAG GTA GGG AGT T
28X	6911	20	CGG ATA CGA CCT TAG AGG CG
rc28X	6930	20	GGC CTC TAA GGT CGT ATC CG
28Z	7200	21	GCA AAG GAT AAG CTT CGA TGG

Degenerate positions are noted by their IUB single-letter code: R = A/G, Y = C/T, M = A/C,K = G/T, W = A/T, H = A/T/C, and N = A/T/C/G.
[a] Sequences for additional less degenerate clade- or taxon-specific primers are available upon request from the senior author.
[b] Amino acid position within an alignment of CAD amino acid residues.
[c] Position of 5′-most position relative to *Drosophila melanogaster* CAD sequence (with introns excised) from genome sequencing project (Adams et al., 2000) or 3′-most position relative to *Drosophila melanogaster* 28S sequence (Hancock et al., 1988).

2.3. DNA sequencing, contig construction, and editing

Purified PCR products were directly sequenced using d-rhodamine Terminator Cycle Sequencing Ready Reaction with AmpliTaqFS DNA Polymerase (PE Applied Biosystems, Foster City, CA) in full reaction volumes, i.e., 8 μl of reaction mix in a 20-μl volume. Comprehensive details about parameters used in CAD sequencing reactions appear in Moulton and Wiegmann (2004). Sequences were gel fractionated and base-called on an ABI PRISM 377 automated DNA sequencer (PE Applied Biosystems, Foster City, CA). Electropherograms from CAD sequencing reactions were trimmed, assembled into contigs, and conflicting calls resolved, where possible, using Sequencher 4.1 (Gene Codes Corp., Ann Arbor, MI) whereas 28S sequences were similarly analyzed using the programs Trev and Gap4 of the Staden package (Bonfield et al., 1995; Staden, 1996).

2.4. Phylogenetic analysis

2.4.1. Nucleotide sequence alignments

Introns within CAD sequences were excised after contig construction using the GT-AG rule (Rogers and Wall, 1980). Alignment of CAD exons was straightforward, except within two regions spanning positions 289–333 and 979–1020 in our alignment. Exclusion of these hypervariable regions had no noticeable effect on topology so they were included in all analyses. Alignment of CAD was performed using Clustal X (Thompson et al., 1997), with slight optimization using Se-Al 2.0 (Rambaut, 2002). The alignment of 28S rDNA sequences was performed manually in GDE 2.2 (Smith et al., 1994), with hypervariable regions in which ad hoc choice of alternative alignments could affect the outcome, removed. The 28S alignment covered the same regions of the gene (D3–D12; Hancock et al., 1988) as were used in Collins and Wiegmann (2002a) and newly obtained sequences in the present study were aligned using Collins and Wiegmann (2002a) as an initial estimate of positional homology among the study taxa. The nucleotide alignments of both genes are available as GDE, NEXUS- or NBRF-formatted files from the authors. The CAD ortholog from *Anopheles gambiae* (GenBank Accession No. EAA06526) was not included in phylogenetic analyses performed here because it created a long-branch attraction problem rendering Cyclorrhapha paraphyletic (Moulton and Wiegmann, 2004). CAD and

28S sequences were tested for homogeneity of nucleotide composition among taxa using a χ^2 test performed in PAUP* 4.0b10 (Swofford, 2002). Only informative sites were included in these χ^2 tests. PAUP* was also used to calculate uncorrected pairwise sequence divergences for all taxa. Codon positions in the CAD data sets were designated using MacClade 4.0 (Maddison and Maddison, 1992).

2.4.2. Character weighting/inclusion

A number of different codon weighting strategies were employed, most of which sought to minimize or eliminate noise from synonymous variation in the third codon position (nt3) data partition relative to largely nonsynonymous variation in the first (nt1) and second (nt2) positions. Strategies employed included equal weighting of all codon positions (equal weights in parsimony analysis; inclusion of all sites in maximum likelihood (ML) and Bayesian analyses), increased weighting of nt2 relative to nt1 and nt3 (codon weighting parsimony, CWP), and total elimination of nt3 (1/1/0 CWP; exclusion of nt3 in ML and Bayesian analyses). When nt3 were excluded in ML and Bayesian analyses, they were also excluded when nucleotide substitution models were assessed using ModelTest (see below).

2.4.3. Parsimony analyses

We conducted phylogenetic analyses of CAD and 28S nucleotides using the parsimony criterion implemented in PAUP* 4.0b10 (Swofford, 2002). All characters were treated as unordered. Trees were constructed by heuristic search with tree bisection-reconnection (TBR) branch swapping in a random stepwise addition of taxa repeated 1000 times. Maxtrees was set to increase incrementally. Node support was evaluated by nonparametric bootstrap resampling (Felsenstein, 1985) and through calculation of Bremer support (Bremer, 1988, 1994). Bootstrap scores were calculated from 1000 replicates, with each replicate consisting of a 10 searches starting with a tree built by stepwise addition using the simple addition sequence. Bremer support was calculated using TreeRot, vers. 2 (Sorenson, 1999).

2.4.4. Maximum likelihood analyses

Phylogenetic trees were also estimated under the maximum likelihood (ML) criterion implemented in PAUP* 4.0b10 (Swofford, 2002). Nucleotide substitution models for each molecular data set were selected using Modeltest 3.06 (Posada and Crandall, 1998). Once a model was selected for a given molecular data set, we used this model and its parameter estimates to search for an optimal ML tree via heuristic searches of tree space using TBR branch swapping in a random stepwise addition of taxa repeated 25 times.

2.4.5. Bayesian analyses

Bayesian Markov Chain Monte Carlo phylogenetic analysis was conducted using primarily MrBayes 2.01 (Huelsenbeck and Ronquist, 2001) using the models and parameters suggested for each data matrix by Modeltest. A mixed models analysis was performed on the entire CAD data set using MrBayes 3.1 (Ronquist and Huelsenbeck, 2003) in an attempt to compensate for extreme nucleotide bias observed at the third codon position between a single empid (*Ceratomerus*) and several outgroup taxa. Each Markov chain in the Bayesian search was started from a random tree and run for one million cycles, sampling every 1000th cycle from the chain. Four chains were run simultaneously, three hot and one cold. Each simulation was run twice. We used the default settings for the priors on the rate matrix (0–100), branch lengths (0–10), and proportion of invariant sites (0–1). Stationarity was evaluated by monitoring likelihood values graphically. The initial 1000 trees in each run were discarded as burn-in samples. The remaining trees were used to construct majority rule consensus trees. Bayesian posterior probabilities for each clade were derived from the trees remaining after discarding the burn-in samples. For ease of visual comparison to bootstrap values, we present these probabilities as whole numbers ranging from 0 to 100. Posterior probabilities greater than or equal to 95% are generally regarded as strong support for a clade (Wilcox et al., 2002).

3. Results

3.1. Data characteristics

3.1.1. CAD

The CAD data set included sequences from 73 dipteran taxa and contained 3915 characters after removal of primer sequences (Table 3). Sequences ranged in size from 3816 (*Leptopeza*) to 3864 (*Bicellaria*) nucleotides. CAD sequences for 43 taxa are presented here for the first time, while 30 are included from a previous study (Moulton and Wiegmann, 2004), including the *D. melanogaster* ortholog obtained from GenBank (Accession No. AE003503). The CAD fragment corresponding to the region between primers 1126F and 1436R could not be obtained for *Chelipoda* sp. for an unresolved reason. Indels were largely confined to the two aforementioned areas of difficult alignment (positions 289–333 and 979–1020).

Base composition in CAD was slightly A/T biased (56%) and differed significantly among taxa ($P < 0.001$). Significant heterogeneity was observed at the first codon position (nt1) and the third codon position (nt3) ($P < 0.001$), but not the second (nt2) ($P < 1.00$). Only three nt3 comparisons, *Rhamphomyia* sp. 38–*Rhamphomyia* "Lake", *Rhamphomyia–Empis*, and *Dolichopus–Neurigona*, did not significantly differ.

Numerous sites in the CAD data set possessed discrepancies between two nucleotides and were coded as ambiguous using the appropriate IUB symbols. Each ambiguous site was double-checked on corresponding electropherograms, and none could be attributed to recognizable sequencing artifacts. All but one (99%) of these ambiguities

Table 3
Descriptive statistics for separate and combined data sets

Data	CAD[a]	28S	CAD[a] + 28S	CAD[b] + 28S
Total length (No. sites analyzed)	3915 (2610)	1342 (1342)	5257 (3952)	5257 (5257)
No. variable sites (%)	826 (43)	389 (29)	1440 (36)	2699 (51)
No. parsimony-informative sites (%)	826 (32)	249 (19)	1005 (25)	2253 (43)
Average nucleotide frequencies				
A	0.31	0.30	0.30	0.32
C	0.19	0.19	0.19	0.17
G	0.25	0.23	0.25	0.22
T	0.25	0.28	0.26	0.29

[a] Third codon positions excluded.
[b] Third codon positions included.

represented synonymous mutations. Given that these sequences were generated by PCR from genomic DNA, and not by cloning, these polymorphisms likely correspond to allelic differences. The total number of base differences between putative alleles ranged from 0 (several taxa) to 27 (*Musca domestica*).

As expected for third positions, substitutions accumulated much faster than at nt1 and nt2. Uncorrected pairwise divergences for nt1, nt2, and nt3 ranged from 4.6% (*Empis*: *Rhamphomyia*) to 26% (*Bombylius*: *Drosophila*), 1.2% (*Empis*: *Rhamphomyia*) to 14% (*Drosophila*: *Opetia*), and 24.3% (*Empis*: *Rhamphomyia*) to 82% (*Bombylius*: *Paraplatypeza*), respectively.

3.1.2. 28S

The 28S data set included sequences from 71 dipteran taxa and was comprised of 1342 nt sites. Sequences for 32 of these taxa are included from a previous study (Collins and Wiegmann, 2002a) of Empidoidea using 28S, while 39 taxa are presented herein for the first time. The ortholog from *D. melanogaster* was obtained from GenBank (Accession No. AE003503). 28S sequences from *Chelipoda, Apalocnemis* "Chile" and *Bicellaria* are truncated at the 5′ end and those of *Lonchoptera*, *Rhamphomyia* sp. 38, *Trichoclinocera*, *Brachystoma* "Chile", and *Clinocera* (CA1) are truncated at the 3′ end. More details about many of the 28S sequences analyzed herein can be found in Collins and Wiegmann (2002a).

Divergence among 28S sequences ranged from 0.08% for comparisons within the genus *Hilara* (Empididae) to 9.6% between outgroup taxa, *Efflatouniella* sp. (Therevidae) and *Phora* sp. (Phoridae). Most pairwise divergences ranged between 1 and 5%.

3.1.3. CAD + 28S

The total combined data set for the two genes included 61 taxa and was 5257 nt in length. With third positions in CAD excluded, the data set is comprised of 3952 characters. Other characteristics of this data set are presented in Tables 3 and 4.

3.2. Phylogenetic inference

3.2.1. CAD

Maximum parsimony analysis of nt1 + nt2 recovered a single island (Maddison, 1991) of two most parsimonious trees

Table 4
Parameters for the models used in maximum likelihood analyses

Partition	CAD nt1 + nt2	28S	CAD (all) + 28S
Gamma distribution shape parameter (α)[a]	0.7131	0.6350	0.4262
Proportion of invariable sites (ϕ)[a]	0.4901	0.5523	0.5875
Relative substitution rate parameters[b]			
A–C	3.2647	1.0779	4.9599
A–G	4.2844	4.3834	8.1154
A–T	1.6959	2.2627	2.6621
C–G	1.5589	0.5717	3.3927
C–T	6.441	4.3834	16.2924
G–T	1.000	1.000	1.000
Estimated nucleotide proportions[a]			
A	0.3099	0.3174	0.3589
C	0.2564	0.1817	0.1411
G	0.2176	0.2195	0.1970
T	0.2161	0.2814	0.3030

The general time reversible (GTR) model was used in analyses of CAD and CAD + 28S sequences. The transversion (TVM) model was used in the analysis of 28S sequences.
[a] α, ϕ, and nucleotide proportions were estimated simultaneously.
[b] R-matrix values.

(MPTs) of 8229 steps (RI = 0.45, RC = 0.11). These two trees differ only with respect to position of *Ceratomerus* + *Oreogeton* relative to a clade comprised of *Apalocnemis*, *Heterophlebus*, and *Hyperperacera* and another clade comprised of the Hemerodromiinae + Empidini. The MPTs differ from the ML tree as follows: (1) *Lonchoptera* and *Opetia* appear as sister taxa, (2) *Pipunculus* is included within Cyclorrhapha, (3) *Acarteroptera* and *Meghyperus occidens* are juxtaposed, (4) *Nematoproctus* is the basal dolichopodid, and (5) the clinocerines + *Hesperempis* form the sister group to a clade comprised of *Apalocnemis*, *Hyperperacera*, and *Heterophlebus*. Results of analyses in which third positions were successively downweighted relative to first and second positions resulted in similar trees differing primarily in relationships among the aforementioned groups. The GTR + I + G model, with percentage of invariant sites and gamma distribution shape parameters estimated to be 0.4984 and 0.7133, respectively, was selected as the most appropriate model by likelihood ratio tests performed within Modeltest (Table 4). Maximum likelihood analysis of CAD nt1 + nt2 using the aforementioned model and parameters and Bayesian likelihood analysis in which these parameters were estimated resulted in similar trees that were largely concordant with the MPTs recovered from these data, differing only with respect to the placement of the empidids *Ceratomerus* sp. and *Oreogeton scopifer*. The maximum likelihood tree ($-\ln L = 39{,}988.56$; MP tree-length = 8255) is depicted as Fig. 1. The Bayesian ($-\ln L = 40{,}146.69$; MP length = 8260) is not shown.

3.2.2. 28S

Maximum parsimony analysis recovered 1153 most parsimonious trees of 1063 steps (CI = 0.488; RI = 0.672). The majority rule consensus of these trees is similar to but less resolved than the trees recovered in maximum likelihood ($-\ln L = 7977.43$; MP length = 1071) and Bayesian likelihood analyses ($-\ln L = 8081.82$; MP length = 1154) of these data. The tree recovered in the Bayesian analysis is shown as Fig. 2.

3.2.3. CAD + 28S

Parsimony analysis of CAD nt1 + nt2 combined with 28S sequences recovered two islands each with a single minimum-length tree (not shown) of 7493 steps (RI = 0.47, RC = 0.14). One of these trees is identical to the MPT from CAD nt1 + nt2 alone, while the other differs from all other recovered topologies in that *Trichopeza* + *Heleodromia* joins *Ceratomerus* and *Oreogeton* in a clade that forms the sister group to the Clinocerinae. Bayesian likelihood analysis of CAD nt1 + nt2 combined with 28S sequences resulted in a tree (not shown, $-\ln L = 41{,}852.16$; MP length = 7557) very similar to the one recovered using maximum likelihood criteria ($-\ln L = 42{,}213.90$; MP length = 7917). These likelihood inferences differ only with respect to the placement of *O. scopifer*.

Parsimony and likelihood analyses of entire CAD (nt1 + nt2 + nt3) and 28S sequences yielded trees (not shown) in which Empidoidea was not monophyletic due to the empid *Ceratomerus* sp. erroneously nesting deeply within the Cyclorrhapha. This arrangement was recovered under each evolutionary model selected, including a mixed models analysis implemented in Mr. Bayes 3.1 (Ronquist and Huelsenbeck, 2003) allowing for different evolutionary rates for separate partitions, in this case nt1 + nt2 versus nt3. Analysis of base composition at the third codon position in the CAD partition revealed a marked increase in cytosine and guanine (C = 0.23242, G = 0.30144) within this empid relative to all other empidoid taxa examined (C: range = 0.13–0.19, mean = 0.15; G: range 0.19–0.24, mean = 0.21). This strong bias in 3rd position base composition is also seen in the cyclorrhaphan taxa included as outgroups and likely explains the clearly convergent grouping of *Ceratomerus* sp. with *Opetia nigra* (C = 0.25402, G = 0.29486), *D. melanogaster* (C = 0.26870, G = 0.29372), and *Paraplatypeza atra* (C = 0.30472, G = 0.29775).

4. Discussion

Nuclear protein-encoding genes are increasing in importance as phylogenetic markers (Mardulyn and Cameron, 1999; Moulton and Wiegmann, 2004; Regier et al., 2004, 2005). These genes are generally more difficult to amplify widely across diverse taxonomic groups, but once obtained they provide a suite of character changes from within the coding portions of the gene that can provide important node support within a phylogenetic analysis. This new character source is particularly valuable for resolving relationships within diverse Mesozoic-aged insect groups like the Empidoidea for which morphological and species diversity has obscured support for phylogeny-based classifications. Our previous study of the carbamoyl phosphate synthase domain of the CAD protein locus in Eremoneuran flies (=Empidoidea + Cyclorrhapha) indicated abundant nucleotide support for divergences at the deeper levels of this large clade, approximately 120–190 mya—a time range that spans the ages of origin of the key lineages comprising empidoid diversity (Wiegmann et al., 2000).

The present work broadens our sample of CAD and 28S ribosomal DNA sequences published in these previous studies for empidoid taxa (Collins and Wiegmann, 2002a,b; Moulton and Wiegmann, 2004). The addition of key taxa and a large amount of gene sequence data greatly increases phylogenetic node support for basal relationships and our revised phylogenetic estimates are largely congruent with our previous findings. Analyses of these larger data sets by parsimony, maximum likelihood and Bayesian methods generally agree in topology and show dramatically improved levels of support for key monophyletic lineages. Although our results are corroborated across analysis strategies and data partitions, we somewhat arbitrarily chose the maximum likelihood tree (Fig. 1) inferred from nt1 + nt2 of the more stringently sampled CAD data set as the best current estimate of empidoid relationships.

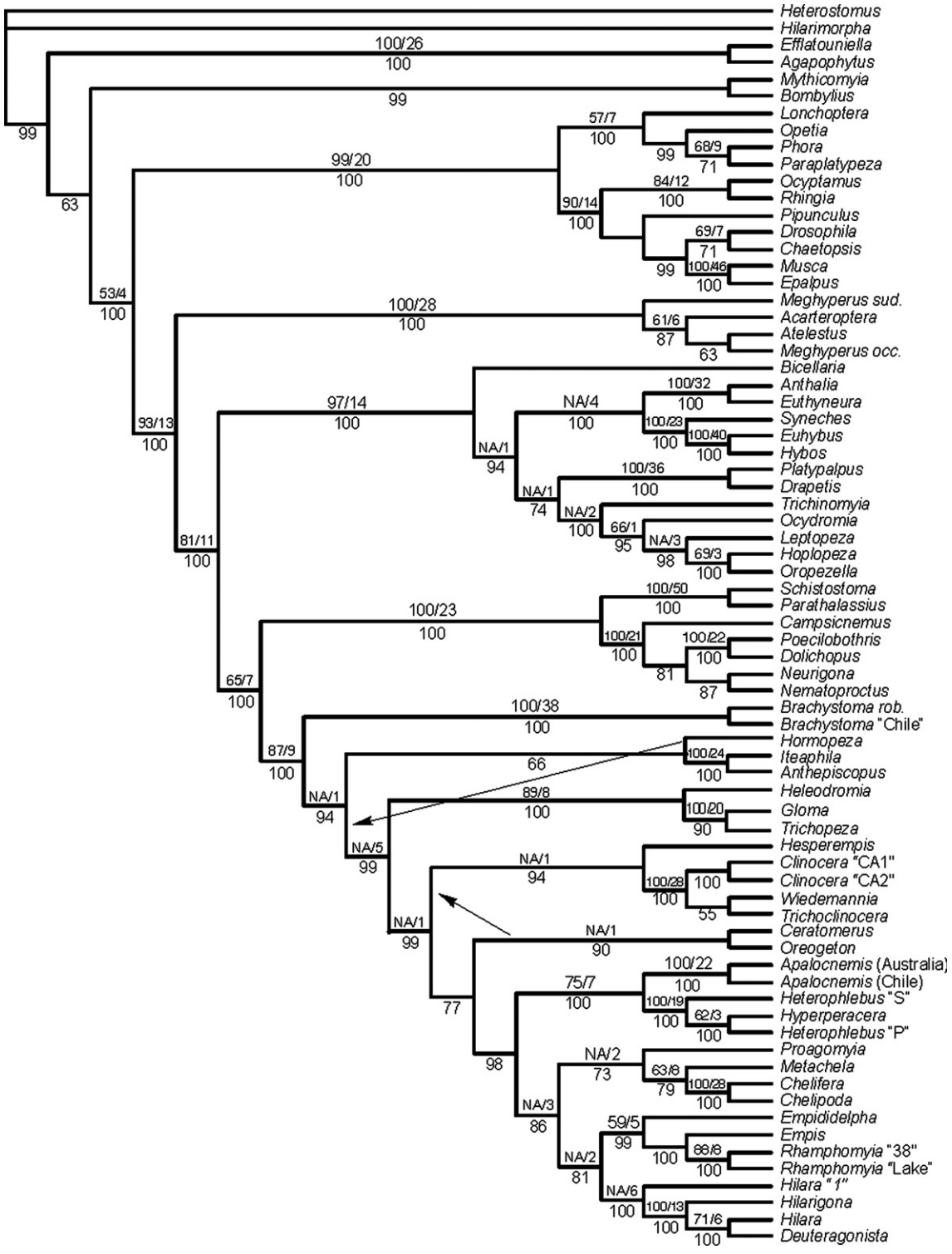

Fig. 1. Phylogenetic tree inferred from first and second codon positions of CAD sequences using maximum likelihood criteria ($-\ln L = 39{,}988.56$; MP treelength = 8255 steps). Model parameters included general-time-reversible (GTR) substitutions (A–C = 3.00; A–G = 3.74; A–T = 1.58; C–G = 1.40; C–T = 5.76; G–T = 1.00), empirical base frequencies (A = 0.302; C = 0.266; G = 0.226; T = 0.207), invariable sites (I = 0.4984), and rate heterogeneity shaped according to a random distribution ($\alpha = 0.7133$). MP bootstrap scores and Bremer support are depicted above nodes (unless bootstrap score is less than 50% and Bremer support is lacking) and Bayesian posterior probabilities are depicted below nodes. Arrows reflect topological discrepancies with the ML tree inferred from concatenated CAD nt1 + nt2 and 28S rDNA sequences.

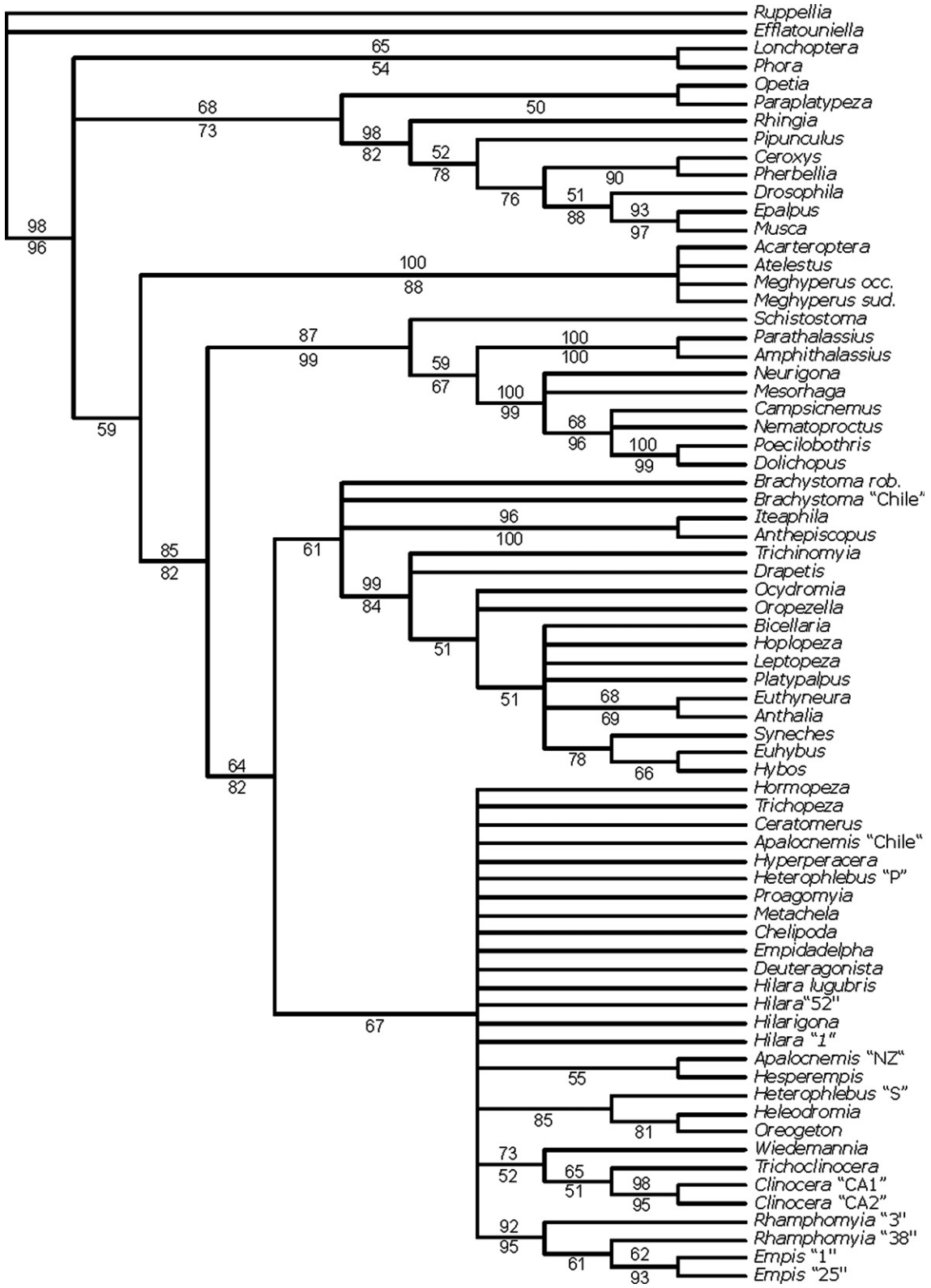

Fig. 2. Phylogeny inferred from 28S rDNA sequences using maximum likelihood criteria ($-\ln L = 7799.2661$; MP treelength = 1041 steps). Model parameters included general-time-reversible (GTR) substitutions (A–C = 1.1182; A–G = 4.3769; A–T = 2.1986; C–G = 0.5759; C–T = 4.3769; G–T = 1.0000), empirical base frequencies (A = 0.31550; C = 0.17920; G = 0.21960; T = 0.28570), invariable sites (I = 0.5503), and rate heterogeneity shaped according to a random distribution ($\alpha = 0.6168$). MP bootstrap scores and Bayesian posterior probabilities are depicted above and below nodes, respectively.

We use this tree to focus further discussion of our phylogenetic findings.

4.1. Higher-level relationships within the Empidoidea

Analyses of CAD nt1 + nt2 alone or concatenated with 28S sequences yielded trees with similar and largely stable topologies that support the five-family empidoidean classification of Chvála (1983) and to a large extent his morphology-based hypotheses about their interrelationships. These five families represent four major empidoid lineages: Atelestidae, Microphorinae + Dolichopodidae, Hybotidae (Tachydromiinae, Ocydromiinae, Hybotinae), and Empididae *s.str*. (all remaining empidoids). A sister group relationship between Atelestidae and the remaining empidoids was recovered with robust support (82–94% bootstrap scores, 100% posterior probabilities) in all analyses conducted. Collins and Wiegmann (2002a) also recovered this arrangement from a 28S rDNA data matrix comprising a more limited sampling of empidoidean taxa. Given the increasingly strong support for a basal position of Atelestidae among the Empidoidea, the morphological similarities in Atelestidae and basal Cyclorrhapha that were emphasized by Chvála, and further amplified in Wiegmann et al. (1993) (but see Cumming et al., 1995 and Sinclair and Cumming, 2006), could be re-interpreted as an early derivation from the ground plan condition of the common ancestor of all Empidoidea (Collins and Wiegmann, 2002a).

Relationships among the three remaining major lineages were also consistently recovered in our analyses and were as follows: Hybotidae, Microphorinae + Dolichopodidae, and Empididae *s.str*. A close relationship between Microphorinae + Dolichopodidae and Hybotidae has been previously proposed based upon morphological (Chvála, 1983) and molecular evidence (Moulton and Wiegmann, 2004). The morphological evidence for this potential grouping consists of the following characters: male terminalia asymmetrical, gonopods fused, and radius 4 + 5 unbranched (Chvála, 1983). However, all of these characters have been found to contain various degrees of homoplasy, or have been questioned as to their homology and scoring (Sinclair and Cumming, 2006; Wiegmann et al., 1993; Cumming et al., 1995). Molecular support for a sister relationship between Microphorinae + Dolichopodidae and Hybotidae was observed in distance analyses of CAD or analyses in which silent variation in third positions was included (Moulton and Wiegmann, 2004). The considerably stronger support observed for a sister relationship between Microphorinae + Dolichopodidae and Empididae *s.str.* found in analyses of CAD nt1 + nt2 and Bayesian analysis of concatenated 28S and CAD sequences (with high posterior probability) compels us to consider this to be the most probable resolution given all available data. Similarly, the most recent morphological study suggests that the Microphorinae + Dolichopodidae are sister to the newly recognized family Brachystomatidae—a group of previously recognized small subfamilies of Empididae (Sinclair and Cumming, 2006). Our data do not corroborate monophyly for Brachystomatidae (Figs. 1 and 2).

A close relationship between Microphorinae and Dolichopodidae was robust in all analyses (99–100% bootstrap, 100% posterior probability) and agrees with results from several major studies of empidoid relationships based on morphology and molecular data (Chvála, 1983; Wiegmann et al., 1993; Cumming et al., 1995; Collins and Wiegmann, 2002a).

4.2. Relationships within families of Empidoidea

4.2.1. Atelestidae

Relationships within Atelestidae were not convincingly resolved due to problems in securing the position of the root within the family, which is shared by the Palearctic species *Atelestus pulicarius* or *Meghyperus sudeticus* in all analyses. When third positions of CAD are included in analyses, a close relationship between the Californian species *Meghyperus occidens* and the Chilean *Acarteroptera recta* is observed with robust support. On the basis of this observation and morphological similarities shared by these two taxa, *Meghyperus* as currently defined is likely polyphyletic (Wiegmann et al., 1993). As a result, we suggest that Nearctic members of *Meghyperus* be transferred to the genus *Acarteroptera* to better define these genera.

4.2.2. Hybotidae

The monophyly of this relatively newly recognized dipteran family is supported in all recent treatments of the superfamily and is strongly re-confirmed here. Our CAD data support a phylogenetic arrangement within the Hybotidae that implies paraphyly for Chvála's Ocydromiinae, an observation that is also supported in most morphological treatments (Cumming et al., 1995; Sinclair and Cumming, 2000, 2006). However, our nucleotide data, whether analyzed separately or combined, indicate robust support for the Tachydromiinae (*Drapetis* + *Platypalpus*) and Hybotinae (*Euhybus*, *Hybos*, *Syneches*). Several additional groupings were observed with robust support (at least in Bayesian terms), including a clade comprised of *Trichinomyia*, *Leptopeza*, *Hoplopeza*, *Ocydromia*, and *Oropezella* and another one containing *Anthalia* and *Euthyneura*. The placement of *Bicellaria* at the base of the Hybotidae (as shown in Fig. 1) was inconsistent among the various analyses of data partitions and methodologies employed, suggesting that increased sampling of both genes and taxa will be required to further resolve relationships within the Hybotidae.

4.2.3. Microphorinae + Dolichopodidae

The monophyly of this clade is well supported, however, we observed mixed support for relationships between the microphorine and dolichopodid genera sampled. The 28S data set provided weak evidence for a sister-group

relationship between the genus *Parathalassius* (Parathalassiinae) and all Dolichopodidae as was first suggested by Hennig (1971) and then echoed in detailed morphological assessments by Chvála (1983), Ulrich (1990), and Sinclair and Cumming (2006). Our CAD data, however, strongly supports (100% bootstrap and posterior probability) the pairing of *Parathalassius* with *Schistostoma*, the only other sampled microphorine, within a monophyletic Microphorinae. Morphological data supports the former hypothesis—the placement of microphorine taxa in a paraphyletic grade of lineages at the base of the Dolichopodidae. Two few microphorine exemplars are included here to add convincing evidence from DNA sequences to support Chvála's view of 'Microphoridae' as a family; increased sampling of microphorines and diverse dolichopodid lineages for CAD and additional protein coding genes is an important goal for future empidoid research.

4.2.4. Empididae sensu stricto

All of our analyses support placement of the remaining sampled empidoid taxa into a monophyletic family Empididae (*s.str.*). The CAD and 28S genes analyzed here provide some modest support at lower taxonomic levels, but full elucidation of relationships within major empidid lineages will require more thorough taxon sampling and use of genes with more variability among genera. One interesting observation from CAD inferences is the basal placement of *Brachystoma*, *Hormopeza*, *Iteaphila*, and *Anthepiscopus.* This phenomenon occurred under various combinations and using differing analysis conditions, although the most parsimonious tree inferred from 28S sequences showed three of these taxa, *Brachystoma*, *Iteaphila*, and *Anthepiscopus*, at the base of the Hybotidae. A slightly different rooting of this 28S inference with respect to the Hybotidae, i.e., one that renders it monophyletic, would show these three empid genera as basal within the Empididae. Interestingly, *Hormopeza*, *Iteaphila*, and *Anthepiscopus* have been placed in the Oreogetoninae (Chvála, 1983), a group considered by Cumming et al. (1995) and Sinclair and Cumming (2006) to either be basal lineages or of uncertain placement within Empidoidea and representative of the plesiomorphic condition for empidoid adult morphology, especially complex features of the male genitalia. Our molecular analyses instead indicate polyphyly for the diverse genera placed in Oreogetoninae in morphological classifications, suggesting to us that oreogetonine morphological scorings assigned to the empidoid groundplan by Cumming et al. may in fact be more parsimoniously interpreted as groundplan features of the Empididae *s.str*.

The Clinocerinae, as currently classified, was not recovered as monophyletic in any analysis performed. *Clinocera*, *Wiedemannia*, and *Trichoclinocera*, however, consistently appeared as a robustly supported clade in analyses of both genes. The other clinocerine clade sampled, *Proagomyia*, consistently appeared as the sister group to the Hemerodromiinae in all analyses containing CAD sequences. Two of the sampled oreogetonines, *Hesperempis* and *Oreogeton*, along with *Ceratomerus*, consistently appeared adjacent or as a sister group to the Clinocerinae in CAD inferences, although only *Hesperempis* appeared as the sister to the aforementioned recovered clinocerine subclade with relevant support (e.g., post. prob. of 94 in CAD nt1 + nt2 Bayesian analysis).

As predicted by Collins and Wiegmann (2002a), the Hemerodromiinae are supported as close relatives of the Empidinae. The sampled hemerodromiine taxa are placed as the sister to the Empidinae in analyses of CAD with Empidinae strongly supported in all analyses containing CAD sequences.

Some support was observed in CAD analyses for close relationships among taxa placed in the Trichopezinae *sensu* Vaillant or Sinclair in that three of the five sampled trichopezine genera consistently form a monophyletic group with strong support, e.g., *Heleodromia* + *Trichopeza* + *Gloma* (Fig. 1). The remaining sampled trichopezines *Apalocnemis* + (*Hyperperacera* and *Heterophlebus*) also form a clade, but are placed in a more derived position within the empidoid tree. Recognition of the monotypic genus *Hyperperacera* appears totally unjustified based upon molecular evidence found in CAD. It should be subsumed into *Heterophlebus*.

4.3. Taxonomic implications and conclusions

The five-family system first proposed by Chvála (1983) for the Empidoidea (Atelestidae, 'Microphoridae' + Dolichopodidae, Hybotidae, Empididae) is increasingly supported by the weight of molecular evidence. We see little reason to delay widespread adoption of this arrangement, with the exception of Microphorinae—the monophyly of which remains difficult to defend on both morphological and molecular grounds. The strong support for these lineages by the current molecular data now provides a robust phylogenetic framework on which new morphological definitions can be based. Subfamily definitions within Empididae *s.str.* remain less certain and relationships within the family will clearly require major revision using both molecular and newly interpreted morphological evidence. The current working concepts of Empidinae, Brachystomatinae, Clinocerinae, and Hemerodromiinae are only partially supported by our molecular evidence, and, as presently defined, the groups 'Oreogetoninae' and 'Trichopezinae' are polyphyletic. Clearly, the great diversity of empidoids has hindered widespread resolution of their relationships by modern phylogenetic methods, but molecular data, analyzed independent of specific morphological interpretations, provides an important source of characters against which specific morphological transformations might be assessed. Chvála's (1983) groundbreaking monograph motivated over 22 years of intensive morphological and molecular phylogenetic research that has focused on testing and further resolving empidoid classification, we fully expect that additional molecular evidence will continue to play a major role in the next stages of this ongoing quest.

Acknowledgments

We thank P. Chandler, K. Collins, J. Cumming, M. Irwin, J. MacDonald, J. Olejnicek, J. Skevington, B. Sinclair, H. Ulrich, D. Webb, and D. Yeates for contributing specimens and assistance with taxonomic identifications. J. Cumming and N.E. Woodley graciously helped in identifying the sampled Chilean specimens of *Brachystoma*. This project was supported by a grant from the US National Science Foundation DEB-0089745 to B.M. Wiegmann and J. Thorne.

References

Bonfield, J.K., Smith, K.F., Staden, R., 1995. A new DNA sequence assembly program. Nucleic Acids Res. 24, 4992–4999.

Bremer, K., 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42, 795–803.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., Rutter, W.J., 1979. Isolation of biologically active ribonucleic acid from sources rich in ribonuclease. Biochemistry 18, 5294.

Chvála, M., 1983. The Empidoidea (Diptera) of Fennoscandia and Denmark. II. General part. The families Hybotidae, Atelestidae, and Microphoridae. Fauna Entomol. Scand. 12, 1–297.

Collins, K.P., Wiegmann, B.M., 2002a. Phylogenetic relationships and placement of the Empidoidea (Diptera: Brachycera) based upon 28S rDNA and EF-1a sequences. Insect Syst. Evol. 33, 421–444.

Collins, K.P., Wiegmann, B.M., 2002b. Phylogenetic relationships of the lower Cyclorrhapha (Diptera: Brachycera) based upon 28S rDNA sequences. Insect Syst. Evol. 33, 445–456.

Cumming, J.M., Sinclair, B.J., Wood, D.M., 1995. Homology and phylogenetic implications of male genitalia in Diptera–Eremoneura. Entomol. Scand. 26, 121–151.

Disney, R.H.L., 1994. Continuing the debate relating to the phylogenetic reconstruction of the Phoridae (Diptera). G. Ital. Entomol. 7, 103–117.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Griffiths, G.C.D., 1972. The phylogenetic classification of Diptera Cyclorrhapha, with special reference to the structure of the male postabdomen. Series Entomol. 8, 1–340.

Hancock, J.M., Tautz, D., Dover, G.A., 1988. Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. Mol. Biol. Evol. 5, 393–414.

Hennig, W., 1971. Insektfossilien aus der unteren Kreide. III. Empidiformia (''Microphorinae'') aus der unteren Kreide und aus dem Baltischen Bernstein; ein Vertreter der Cyclorrhapha aus der unteren Kreide. Stuttg. Beitr. Naturkd. 232, 1–28.

Huelsenbeck, J.P., Ronquist, F.R., 2001. MrBayes 2: Bayesian inference of phylogeny. Biometrics 17, 754–755.

Maddison, D.R., 1991. The discovery and importance of multiple islands of most-parsimonious trees. Syst. Zool. 40, 315–328.

Maddison, W.P., Maddison, D.R., 1992. MacClade: Analysis of Phylogeny and Character Evolution, version 4.04 for Mac OS X. Sinauer, Sunderland, MA.

Mardulyn, P., Cameron, S.A., 1999. The major opsin in bees (Insecta Hymenoptera): a promising nuclear gene for higher level phylogenetics. Mol. Phylogenet. Evol. 12, 168–176.

Moulton, J.K., Wiegmann, B.M., 2004. Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged eremoneuran Diptera (Insecta). Mol. Phylog. Evol. 31, 363–378.

Posada, D., Crandall, K.A., 1998. ModelTest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rambaut, A., 2002. Sequence alignment editor, Version 2.0. Available as shareware from <http://evolve.zoo.ox.ac.uk/index.html/>.

Regier, J.C., Shultz, J.W., Kambic, R.E., 2004. Phylogeny of basal hexapod lineages and estimates of divergence times. Ann. Entomol. Soc. Am. 97, 411–419.

Regier, J.C., Wilson, H.M., Schultz, J.W., 2005. Phylogenetic analysis of Myriapoda using three nuclear protein-coding genes. Mol. Phylog. Evol. 34, 147–158.

Rogers, J., Wall, R., 1980. A mechanism for RNA splicing. Proc. Natl. Acad. Sci. USA 77, 1877–1879.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sinclair, B.J., 1995. Generic revision of the Clinocerinae (Empididae), and description and phylogenetic relationships of the Trichopezinae, new status (Diptera: Empidoidea). Can. Ent. 127, 665–752.

Sinclair, B.J., Cumming, J.M., 2000. Revision of the genus *Apterodromia* Oldroyd (Diptera: Empidoidea), with a redefinition of the tribe Ocydromiini. Rec. Aust. Mus. 52, 161–186.

Sinclair, B.J., Cumming, J.M., 2006. The morphology, higher-level phylogeny and classification of the Empidoidea (Diptera). Zootaxa 1180, 1–172.

Smith, S.W., Overbeek, R., Woese, C.R., Gilbert, W., Gellevet, P.M., 1994. The genetic data environment and expandable GUI for multiple sequence analysis. Comput. Appl. Biosci. 10, 671–675.

Sorenson, M.D., 1999. TreeRot, version 2.. Boston University, Boston, MA.

Staden, R., 1996. The Staden sequence analysis package. Mol. Biotechnol. 5, 233–241.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Ulrich, H., 1990. Evidence for the phylogenetic position of Parathalassinae (Empidoidea) based on thoracic morphology. Abstract Volume. The Second International Congress of Dipterology, Bratislava. (Full MS published privately by author).

Wiegmann, B.M., Mitter, C., Thompson, F.C., 1993. Evolutionary origin of the Cyclorrhapha (Diptera): tests of alternative morphological hypotheses. Cladistics 9, 41–81.

Wiegmann, B.M., Mitter, C., Regier, J.C., Friedlander, T.P., Wagner, D.M., Nielson, E.S., 2000. Nuclear genes resolve Mesozoic-aged divergences in the insect order Lepidoptera. Mol. Phylog. Evol. 15, 242–259.

Wilcox, T.P.D., Zwickl, J., Heath, T., Hillis, D.M., 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Mol. Phylogenet. Evol. 25, 361–371.

Yang, D., 2004. Fauna Sinica Insecta (vol. 34) Diptera Empididae Hemerodromiinae Hybotinae. Henan Science & Technology Publishing House. 335 pages + 1 plate.

Zatwarnicki, T., 1996. A new reconstruction of the origin of the eremoneuran hypopygium and its implications for classification (Insecta: Diptera). Genus 3, 103–175.