ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Molecular Phylogenetics and Evolution 27 (2003) 45–57

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

# Can the current molecular arsenal adequately track rapid divergence events within Simuliidae (Diptera)?

John K. Moulton*

*Department of Entomology, College of Agriculture and Life Sciences, Forbes Building, The University of Arizona, Tucson, AZ 85721, USA*

Received 21 December 2001; revised 24 September 2002

**Abstract**

Ancient rapid divergence events, such as those that took place during the Mesozoic, are pervasive in evolution and represent a major challenge to phylogenetic biologists. The number of molecular phylogenetic studies in which rapid divergence has been invoked to account for poor phylogenetic resolution has steadily increased over the past few years. In this study, rapid divergence events are again hypothesized to have taken place, this time within the two major tribes of Simuliidae, Prosimuliini and Simuliini. This inference is based upon the failure of portions of 28S rDNA, EF-1α, DDC, PEPCK, and 12S rDNA to adequately reconstruct relationships among their constituent genera and the presence of short internal and long terminal nodes within both tribes for all character partitions of these genes. Sequence divergence, other than synonymous variation within coding genes, was low among genera and node support weak, except largely for those joining morphologically similar taxa previously recognized as closely related. Strong attraction between a long terminal node (*Austrosimulium* Tonnoir) and a long internal node (Simuliini), is hypothesized to be the reason for strong support for the placement of *Austrosimulium* as the basal-most lineage in this tribe. In spite of these problems, a preferred tree intended to be a reasonable estimate of simuliid phylogeny is tentatively presented. Based upon the considerable genomic sampling conducted in this and previous studies, it is clear that new types of genes are needed to more adequately resolve rapid divergence phenomena. The CAD and GART loci, currently under development as phylogenetic markers by the author, show greater promise for resolving simuliid relationships than do any of the genes examined herein.


## 1. Introduction

Mesozoic-aged divergences are a major hurdle for phylogenetic biologists, particularly those who rely upon molecular sequences as data, because few available markers appear to adequately track these events, especially if they occurred rapidly, e.g., as explosive radiations. Molecular phylogenetic studies in which rapid divergence has been invoked to account for poor phylogenetic resolution among Mesozoic-aged taxa (Abouheif et al., 1998; Halanych, 1998; Hickson, 1993; Kraus and Miyamoto, 1991; Orti and Meyer, 1996; Regier and Schultz, 1998; Roger et al., 1999) are numerous and often involve the use of nuclear ribosomal and mitochondrial ribosomal and coding genes as markers. Because these genes are easily obtained using universal primers (Kocher et al., 1989), Caterino et al. (2000), hoping to facilitate phylogenetic comparisons across a wide array of taxa, advocated their use simply for the sake of convenience, in spite of the growing number of studies revealing severe limitations in the competence of many of them as phylogenetic indicators (Blouin et al., 1998; Foster et al., 1997; Halanych and Robinson, 1999; Hasegawa and Hashimoto, 1993; Hwang et al., 1998; Liu and Beckenbach, 1992; Norris et al., 1999; Simon et al., 1994).

Despite the exponential increase in the number of sequences being deposited in molecular databases, discovery and implemention of genes better suited for addressing these types of divergences has been minimal. In the last decade the availability and appropriateness of nuclear protein coding genes for phylogeny reconstruc-

* Present address: Department of Entomology and Plant Pathology, College of Agricultural Sciences and Natural Resources, The University of Tennessee Knoxville, TN 37996-4560, USA.

*E-mail address:* kevin_moulton@ncsu.edu.

1055-7903/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1055-7903(02)00397-4

tion has been extensively addressed (Brower and DeSalle, 1994; Friedlander et al., 1992, 1994, 1996; Graybeal, 1994). Since this time, however, only five genes—phosphoenolpyruvate carboxykinase (Friedlander et al., 1992, 1996), dopa decarboxylase (Fang et al., 1997), elongation factor-1α (Cho et al., 1995; Mitchell et al., 1997), wingless (Brower and DeSalle, 1998), and opsin (Mardulyn and Cameron, 1999)—potentially useful for resolving Mesozoic-aged divergences within insects have emerged. Despite these valiant exploratory efforts, only EF-1α has seen noticeable implementation in insect molecular systematics.

Black flies (Simuliidae) are a prominent dipteran clade dating back to the Mesozoic period. *Simulimima grandis* Kalugina (Kalugina and Kovelev, 1985), known only from a single pupal depression dating back to the middle Jurassic, ca. 160–170 mybp, was assigned to Simuliidae by Crosskey (1981), although some disagree about the validity of this action. Jell and Duncan (1986) reported several simuliid fossils from the Koonwarra fossil beds of southern Victoria, Australia, which date back to the Aptian phase of the lower Cretaceous, ca. 113–119 mybp. Currie (1988) considered only the larval depressions from the Koonwarra beds to be simuliids and noted similarities between them and an unnamed extant Australian simuliine genus. Kalugina (1991) described three simuliids, *Baisomyia incognita*, *Gydarina karabonica*, and *Kovalevimyia lacrimosa*, from fossil adults preserved in Upper Jurassic or Lower Cretaceous rocks from Transbaikalia and affiliated them with *Gymnopais* Stone/*Twinnia* Stone and Jamnback, *Gymnopais* and *Prosimulium*, respectively. The fact that all of these Mesozoic-aged fossil simuliids can arguably be assigned to extant genera within the two major tribes, Prosimuliini and Simuliini, within Simuliidae serves as a testament to the antiquity and morphological homogeneity of the group and places the origins of both taxa to at least the late Cretaceous.

With basal lineages (subfamilies, tribes) within Simuliidae now well resolved from both morphological (Currie, 1988) and molecular data (Moulton, 2000), research can now be focused towards elucidating relationships among genera within two major clades, Prosimuliini and Simuliini, identified by these previous studies. These clades comprise over 99% of the genera and species within the family, with the latter comprising over 75% of these totals. Relationships within Prosimuliini are reasonably well understood due largely to the work of Currie (1988), but those within Simuliini remain tenuous. Rubtzov (1974) and Py-Daniel (1990) are the only workers who have attempted to view simuliine relationships in a phylogenetic framework, but problems inherent in both studies, namely non-numerical analyses and incomplete taxon sampling, respectively, prevent meaningful conclusions to be drawn.

The goal of the research reported herein was to reconstruct Mesozoic-aged relationships within Prosimuliini and Simuliini using molecular sequences from multiple unlinked loci. Nucleotide sequences representing portions of five genes—28S, EF-1α, DDC, PEPCK, and 12S—generally considered to be useful for resolving divergences of varied age, including Mesozoic-aged ones (Cho et al., 1995; Fang et al., 1997; Friedlander et al., 1992, 1994, 1996; Simon et al., 1994; Vossbrinck, 1989), were obtained from several simuliid genera plus an outgroup and subjected to independent and simultaneous phylogenetic analysis. Inferences drawn from these analyses are compared to one another and to previous estimates drawn from morphological and cytological data. Finally, characteristics of genes perhaps better suited to address ancient groups containing rapidly diverged clades are hypothesized.

## 2. Materials and methods

### 2.1. Taxon sampling

The name, life history stage, and geographical source of the specimens sequenced for this study are provided in Table 1. *Andropsopa gillespieae* (Arnaud and Boussy) and *Parasimulium crosskeyi* serve as distal and proximal outgroups, respectively, for each data matrix. Generic limits within Simuliidae follow those in Crosskey's (1988) conspectus of the family, with some modification in light of findings by Currie (1988). Crosskey and Howard (1997) differ most dramatically from Crosskey (1988) in that the Australian "*Cnephia*" or Prosimuliini (Crosskey, 1981, 1988; Mackerras and Mackerras, 1948, 1949, 1950, 1952, 1955; Rubtzov, 1962) is relegated to *Paracnephia* Rubtzov, as is *Procnephia* Crosskey. I agree with the *Paracnephia* + *Procnephia* arrangement, as did Rubtzov (1962), but not with *Paracnephia sensu* Crosskey and Howard (1997). Having seen well-preserved material of all life stages of nearly all of these Australian species, their relegation, especially as a whole, into a single clade is unjustified. As a result, I treat these species in a manner similar to that of Crosskey (1988) and Currie (1988), i.e., assigned to *Cnephia* but obviously belonging elsewhere. Quotation marks, e.g., "*Cnephia*" *strenua*, are used to stress the inappropriateness of some current generic assignments.

More than one species were examined for *Austrosimulium*, *Gigantodax*, *Greniera*, and *Simulium*. This was done in an attempt to shorten potentially long branches (Felsenstein, 1978), test the monophyly of *Austrosimulium*, *Greniera*, and *Simulium*, and determine relationships within the latter. Of the species of *Simulium* examined, *Simulium* (*Hellichiella*) *congareenarum*, *Simulium* (*H.*) *canonicolum*, and *Simulium* (*Nevermannia*) sp.

Table 1
Name, life history stage, and geographical source of material used in this study

| Taxon | Life stage | Origin/source |
|---|---|---|
| THAUMALEIDAE | | |
| *Androprosopa gillespieae* (Arnaud & Boussy) | L | USA: Arizona |
| SIMULIIDAE | | |
| Parasimuliinae[a] | | |
| *Parasimulium crosskeyi* Peterson | M | USA: Oregon |
| Simuliinae[a] | | |
| Prosimuliini[b] | | |
| *Gymnopais fimbriatus* Wood | L | USA: Alaska |
| *Helodon onychodactylus* Dyar & Shannon | L | USA: Arizona |
| *Prosimulium formosum* Shewell | L | USA: Arizona |
| *Prosimulium impostor* Peterson | L | USA: Arizona |
| *Twinnia nova* Dyar & Shannon | L | USA: Utah |
| *Urosimulium aculeatum* Rivosecchi | L | Spain |
| Simuliini[b] | | |
| *Austrosimulium bancrofti* Taylor cpx. | L | Australia: Victoria |
| *Austrosimulium bancrofti* Taylor (Ipswich A1) | ? | GenBank database |
| *Austrosimulium mirabile* Mackerras & Mackerras | L | Australia: Queensland |
| "*Austrosimulium*" *colboi* Davies & Gyorkos | L | Australia: Victoria |
| *Cnephia ornithophilia* Davies, Peterson, & Wood | L | USA: South Carolina |
| "*Cnephia*" *aurantiacum* Tonnoir | L | Australia: Victoria |
| "*Cnephia*" *strenua* Mackerras & Mackerras | P | Australia: Queensland |
| "*Cnephia*" *tonnoiri* Drummond | L | Australia: WA |
| "*Cnephia*" *umbratorum* Tonnoir | F | Australia: Victoria |
| "*Cnephia*" 'x' | P | Australia: WA |
| "*Cnephia*" 'y' | P | Australia: WA |
| "*Cnephia*" sp. nr. *terebrans* Tonnoir (='GKW2') | L | Australia: WA |
| "*Cnephia*" 'S. x. (east)' | L | Australia: Victoria |
| "*Cnephia*" 'S. x. (west)' | L | Australia: WA |
| "*Cnephia*" *pilfreyi* Davies & Gyorkos | L | Australia: WA |
| *Cnesia disimilis* Edwards | L | Argentina |
| *Crozetia crozetensis* Womersley | L | Isle de Crozet |
| *Ectemnia reclusa* Moulton & Adler | L | USA: South Carolina |
| *Greniera denaria* Davies, Peterson, & Wood | L | Canada: Ontario |
| *Greniera fabri* Doby and David | L | Spain |
| *Gigantodax adleri* Moulton | L | USA: Arizona |
| *Gigantodax marginalis* Edwards | L | Argentina |
| *Mayacnephia* sp. nr. *osborni* Stains & Knowlton | L | USA: Arizona |
| *Metacnephia sommermanae* Stone | L | USA: Alaska |
| *Paracnephia thornei* de Meillon | L | Rep. of South Africa |
| *Paraustrosimulium anthracinum* Moore | L | Chile |
| *Procnephia rhodesiense* (Crosskey) | L | Zimbabwe |
| *Simulium* (*Edwardsellum*) *sirbanum* Vajime & Dunbar | L | Liberia |
| *Simulium* (*Hellichiella*) *canonicolum* (Dyar & Shannon) | L | USA: Arizona |
| *Simulium* (*Hellichiella*) *congareenarum* Dyar & Shannon | L | USA: South Carolina |
| *Simulium* (*Hellichiella*) *curriei* Adler and Wood | L | USA: Arizona |
| *Simulium* (*Nevermannia*) *pugetense* Dyar & Shannon cpx. | L | USA: Arizona |
| *Simulium* (*Psilozia*) *encisoi* Vargas & Díaz Nájera | P | USA: Arizona |
| *Simulium* (*Simulium*) *reptans* Linnaeus | F | UK: England |
| *Stegopterna* "W" (Madahar, 1969) | L | USA: Arizona |

[a] *sensu* Crosskey (1988), Currie (1988), and Moulton (2000).
[b] *sensu* Currie (1988) and Moulton (2000).

nr. *pugetense* are representative of the presumedly plesiomorphic, univoltine, ornithophilic assemblage of subgenera and *Simulium* (*Edwardsellum*) *sirbanum* and *Simulium* (*Simulium*) *reptans* are representative of the presumedly more derived, multivoltine, mammalophilic group of subgenera. Inclusion of a representative of *S.* (*Edwardsellum*) is especially important in light of the findings of Tang et al. (1996). Additional issues concerning taxon and genomic sampling in this study are discussed in Moulton (2000).

### 2.2. DNA manipulation

Procedures for DNA extraction, amplification, and sequencing are described elsewhere (Moulton, 2000). GenBank Accession Nos. for these sequences range

from AF007297 to AF007374 for 28S, AF003552 to AF003582 for EF-1α, AF047531 to AF047541 for PEPCK, AF078875 to AF078888 for DDC, and AF049471 to AF049485 for 12S.

Alignment of EF-1α, DDC, and PEPCK sequences was straightforward, whereas that of the ribosomal genes was not. Initial alignment of 28S was done using Clustal W (Thompson et al., 1997) and subsequently optimized manually using SeqApp 1.9 (Gilbert, 1992). Alignment of 12S proved more difficult due to extreme A/T bias. To alleviate effects of alignment upon phylogenetic inferences from 12S, including having to discard much needed variation, the elision method (Wheeler et al., 1995) was used. Ten independent alignments of 12S were generated using Clustal W. The gap opening/gap extension costs were as follows: 2/1, 5/1, 5/2; 5/10, 7/2; 10/2, 10/5; 20/5, 20/10, and 40/1. These alignments were then concatenated using SeqApp. Sequence alignments used in this study are published elsewhere (Moulton, 1997) and are available upon request from the author as NEXUS or NBRF-formatted files.

### 2.3. Sequence analysis

Uncorrected pair-wise distances, calculations of and tests for biases in nucleotide composition, and all phylogenetic inferences, were performed using PAUP*d52 and PAUP*d54 (Swofford, 1996, 1997; used with author's permission). 28S, EF-1α, and 12S nucleotides were subjected to maximum parsimony, minimum evolution, and neighbor joining analysis. MacClade 3.0 (Maddison and Maddison, 1992) was used to designate nucleotide positions within coding sequences and translate them into amino acid sequences, build constraint trees, and analyze molecular character evolution.

PEPCK and DDC nucleotides were subjected to maximum and codon weighting parsimony analyses. PEPCK and DDC amino acid sequences were also subjected to maximum parsimony and distance analyses. Parsimony and minimum evolution searches used PAUP*'s heuristic search option, with TBR branch rearrangement. Characters were treated as unordered in all parsimony analyses and MAXTREES was set to increase incrementally. Most parsimonious trees were found by conducting 1000 searches, with each search begun from trees acquired by stepwise addition with a random addition sequence order. Each island of most parsimonious trees (Maddison, 1991) was found at least 25 times using this approach. Minimum evolution trees were found by conducting a single search from a tree acquired by neighbor-joining. HKY85 (Hasegawa et al., 1985) corrected distances were used in distance analyses of nucleotide sequences and mean distances were used in analyses of amino acid sequences.

Node support was evaluated by bootstrap resampling (all matrices) and through calculation of decay indices (28S and EF-1α only). Bootstrap values for parsimony, minimum evolution, and neighbor-joining analyses were calculated with 500 bootstrap replicates; for parsimony bootstrapping, each replicate consisted of a single search starting with a tree built by stepwise addition using the simple addition sequence, whereas for distance bootstrapping, each replicate was begun from the neighbor-joining tree. A MAXTREES of 200 was imposed on all bootstrap analyses. Decay indices (Bremer, 1988) were calculated for clades exhibiting greater than 50% bootstrap support. Calculations involved constraining PAUP*'s search to find the most parsimonious trees without that group, which was accomplished by forcing the particular group to be non-monophyletic using a constraint tree, and then comparing the lengths of these trees to those obtained from unconstrained parsimony analysis.

Several codon weighting schemes were investigated in attempts to extract phylogenetic signal from non-synonymous variation within DDC and PEPCK. These assigned first, second, and third codon positions the following weights: 1/1/0, 3/5/1, 3/10/1, and 5/20/1. The choice of weights was entirely exploratory; however the downweighting of third positions relative to the others and the higher weight afforded to second positions relative to first positions is logical (Irwin et al., 1991; Li and Graur, 1991; Nei, 1978).

Variation in 28S is minimal within simuliids, e.g., pair-wise distances among prosimuliine and simuliine genera range from 2 to 6%. Synonymous variation within EF-1α is high, with 372 of 413 third codon positions variant and 341 of these phylogenetically informative. Due to the extremely short internal branches of trees inferred from both genes (see Fig. 6), the low probability of slowly evolving loci, such as ribosomal genes, to reliably track rapid divergence events (Hickson, 1993; Kraus and Miyamoto, 1991; Ledje and Arnason, 1996), the saturated nature of the variation within EF-1α (Moulton, 1997, 2000), and the near absence of strongly supported discordance between these genes (see below), a combined analysis of these data seemed appropriate and was conducted.

In an effort to prevent saturated non-synonymous variation within EF-1α from overwhelming signal from 28S in combined analyses, thus potentially "swamping" clades strongly supported by the latter, clades strongly supported by 28S alone were constrained to be monophyletic in simultaneous analyses of these data. Strong support from EF-1α for a sister group relationship between "*Cnephia*" *pilfreyi* and *Ectemnia* contradicted the 28S (ME/NJ analyses only) placement of "*Cnephia*" *pilfreyi*; which was as the sister group to *Paraustrosimulium* and "*Austrosimulium*" *colboi*. Although bootstrap values were greater for the EF-1α arrangement, the 28S topology was favored because convincing morphological evidence supports it (J.K. Moulton, unpublished).

A "total evidence" approach was also conducted. This data matrix included the following taxa and sequences: *A. gillespieae* (all genes), *P. crosskeyi* (all genes), a *Prosimulium* amalgam of *Prosimulium formosum* (28S, EF-1α, PEPCK, and DDC) + *Prosimulim uinta* (12S), *Helodon onychodactylus* (all genes), *Gymnopais fimbriatus* (all genes), *Austrosimulium bancrofti* (all genes), *Paracnephia thornei* (28S, EF-1α, and DDC), *Cnesia dissimilis* (28S, EF-1α, PEPCK, and 12S), *Paraustrosimulium anthracinum* (28S, EF-1α, and 12S), *Greniera fabri* (28S, EF-1α, PEPCK, and DDC), *Metacnephia sommermanae* (28S, EF-1α, PEPCK, and 12S), and a *Simulium* amalgam of *S. congareenarum* (28S, EF-1α, and DDC) + *S. curriei* (12S).

## 3. Results of phylogenetic inference

### 3.1. Independent analyses

*28S.* Maximum parsimony analysis yields one island of 36 trees (TL = 1044, CI = 0.597, RC = 0.368). A strict consensus of these trees is shown as Fig. 1A. The minimum evolution (ME) tree (score = 0.69886) and neighbor joining (NJ) tree differ from the most parsimonious (MP) trees, most notably by the position of the root within Simuliini; a strict consensus of the ME and NJ trees is shown as Fig. 1B.

*EF-1α.* Maximum parsimony analysis of all nucleotides (equal weighting of codons) yields two trees residing on separate islands (TL = 2398, CI = 0.310, RC = 0.143); a strict consensus of these trees is shown as Fig. 2A. The ME (score = 1.78044) and NJ trees differ considerably from one another in several aspects; the NJ tree is shown as Fig. 2B.

*PEPCK.* Maximum parsimony analysis of all nucleotides yields two trees of length 607 (CI = 0.596, RC = 0.266), one of which is identical to the MP tree found using 3/5/1 and 3/10/1 codon weighting schemes. Parsimony analysis of first and second codon positions yields one island containing 20 trees of length 115 (CI = 0.704, RC = 0.367). Application of a 5/20/1 weighting scheme yields a slightly different tree (Fig. 3) (TL = 612, CI = 0.592, RC = 0.257) than others recovered with different codon weighting schemes. The minimum evolution (score = 0.28814) and neighbor-joining trees (not shown) inferred from all nucleotides differ only in the position of one simuliine genus. The

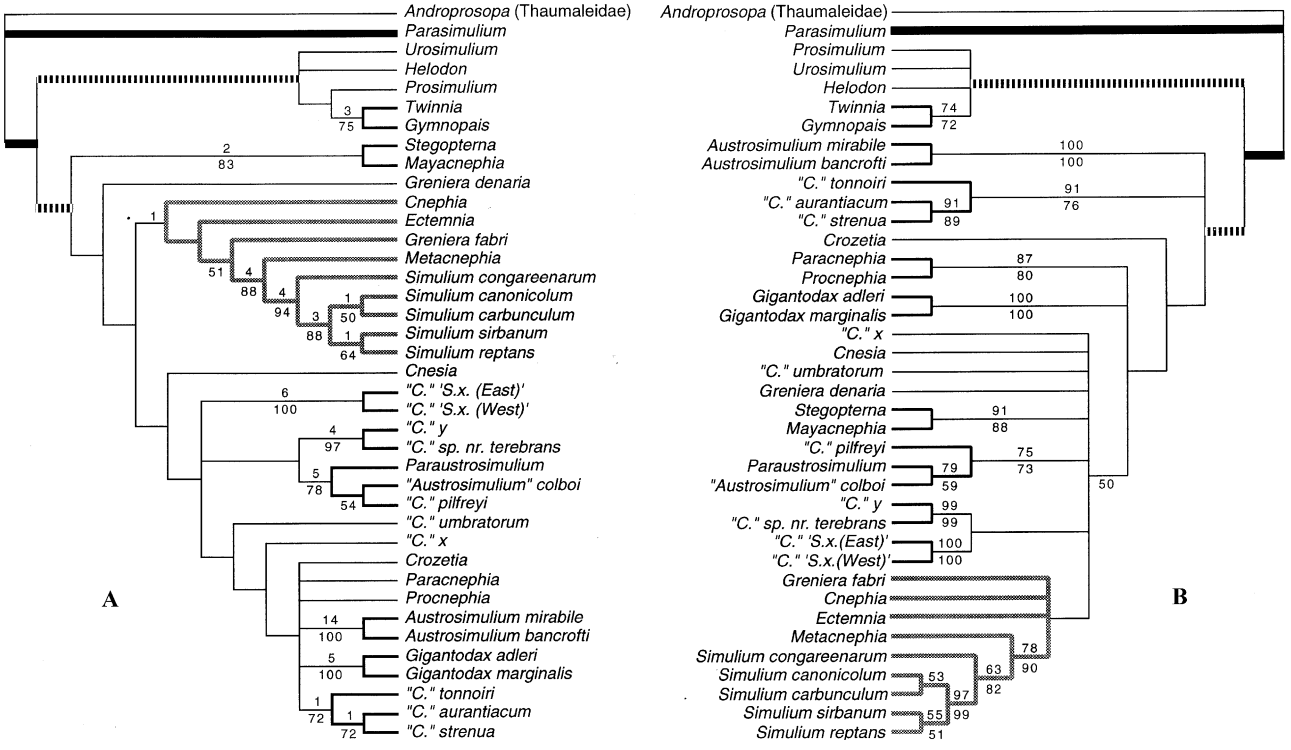


Fig. 1. Results from phylogenetic analysis of 28S sequences. (A) Strict consensus of 36 most parsimonious trees (TL = 1044, CI = 0.597, and RC = 0.368). Decay and bootstrap support for clades appearing in 50% or more of 1000 bootstrap replicates conducted are shown above and below nodes, respectively. (B) Consensus of minimum evolution (score = 0.69886) and neighbor-joining trees. Bootstrap support for clades appearing in 50% or more of 1000 replicates conducted using minimum evolution and neighbor-joining criteria appear above and below nodes, respectively. Subfamilies and tribes are depicted by solid and hatched thick lines, strongly supported intratribal clades are depicted by bold black lines, and major intratribal concordances are depicted by bold gray lines. Subfamily and tribal bootstrap and decay support not shown.

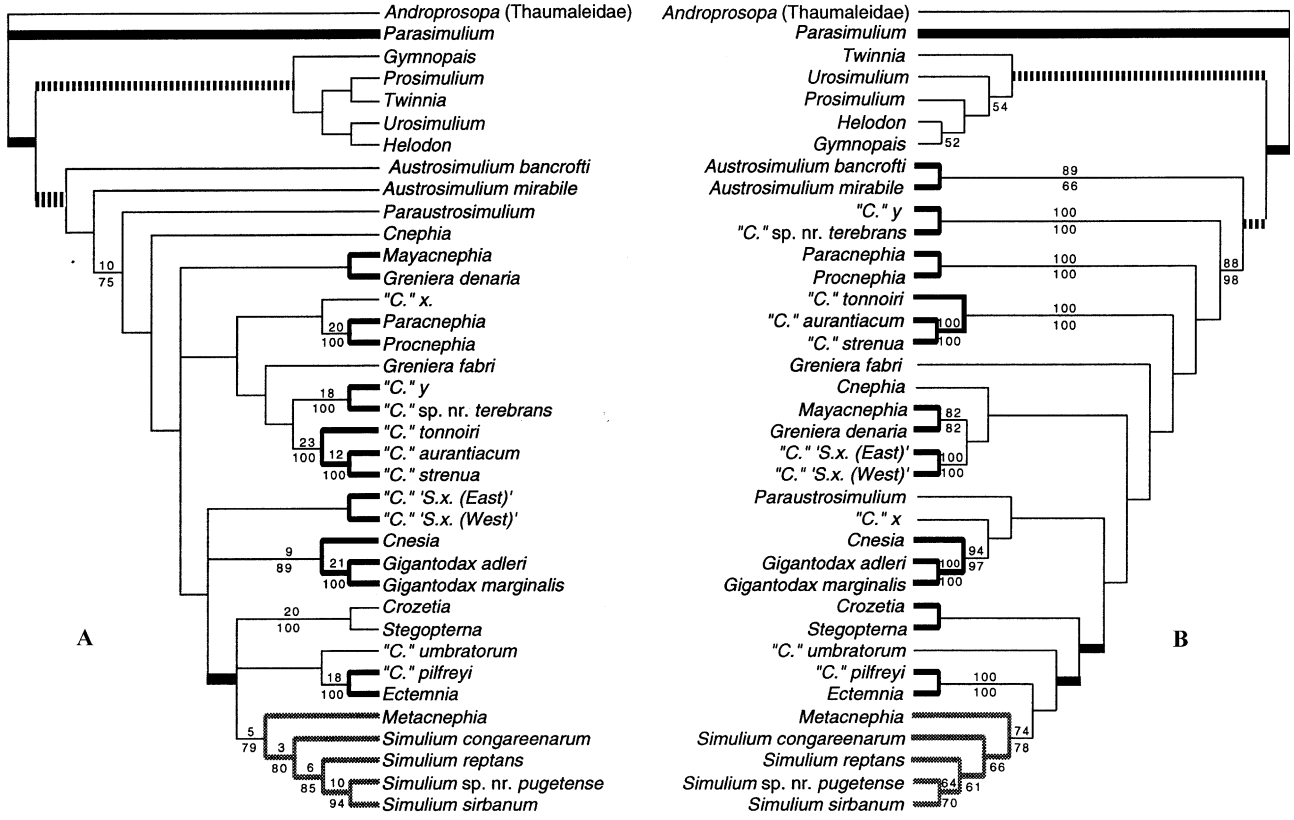


Fig. 2. Results from phylogenetic analysis of EF-1α nucleotide sequences. (A) Consensus of two most parsimonious trees (TL = 2398, CI = 0.310, and RC = 0.143). Decay and bootstrap support for clades appearing in 50% or more of 1000 bootstrap replicates conducted are shown above and below nodes, respectively. (B) Neighbor-joining tree (minimum evolution tree, score = 0.1.78044, not shown). Bootstrap support for clades appearing in 50% or more of 1000 replicates conducted using minimum evolution and neighbor-joining criteria appear above and below nodes, respectively. Subfamilies and tribes are depicted by solid and hatched thick lines, strongly supported intratribal clades are depicted by solid bold lines, and major intratribal concordances are depicted by hatched bold lines. Subfamily and tribal bootstrap and decay support not shown.

minimum evolution (score = 0.19822) and neighbor-joining trees inferred from first and second codon positions (not shown) differ only with respect to the position of the root within Simuliini.

*DDC.* Maximum parsimony analysis of all nucleotides yield two trees of length 663 (CI = 0.575, RC = 0.268). A strict consensus of the trees is shown as Fig. 4A. Maximum parsimony analysis of first and second codon positions yield 36 trees of length 128 (CI = 0.719, RC = 0.465). The minimum evolution tree inferred from all nucleotides (score = 1.24577) is identical to one of the MPTs with respect to relationships among Prosimuliini, except for position of the root (Fig. 4B). The 3/5/1 and 3/10/1 codon weighting searches found the same tree, one of which was identical to one of four MP trees (two islands of 1 and 3 trees) found using 5/20/1 codon weighting (trees not shown). Trees inferred from amino acid sequences (Fig. 4C) were also generally poorly supported and differed from those inferred from nucleotides with respect to simuliine relationships.

*12S.* Maximum parsimony analysis of the elision matrix yields one tree of length 4662 (CI = 0.622, RC = 0.214). The MP tree and ME tree (score = 0.8879) do not contain a monophyletic Simuliini. The NJ tree (Fig. 5) contains monophyletic tribes and shows all of the well-supported intratribal relationships present in the MP and ME trees.

### 3.2. Simultaneous analyses

*28S + EF-1α.* Three MPTs of length 3428 (CI = 0.395, RC = 0.193) representing two islands were found. The MP, ME (score = 1.16784) and NJ trees are quite different; a strict consensus of these five trees results in a topology identical to that of the constraint tree. The ME tree (Fig. 6) is tentatively presented as the most reasonable estimate of simuliid relationships based upon concordance with morphological data and the total evidence tree.

*Total evidence approach.* Parsimony analysis of concatenated 28S, EF-1α, DDC, PEPCK, and 12S nucleotides yields a single tree (TL = 3265, CI = 0.698, and RC = 0.342). The MP, ME (score = 0.66741), and NJ trees differ only in the juxtaposition of *C. dissimilis* and

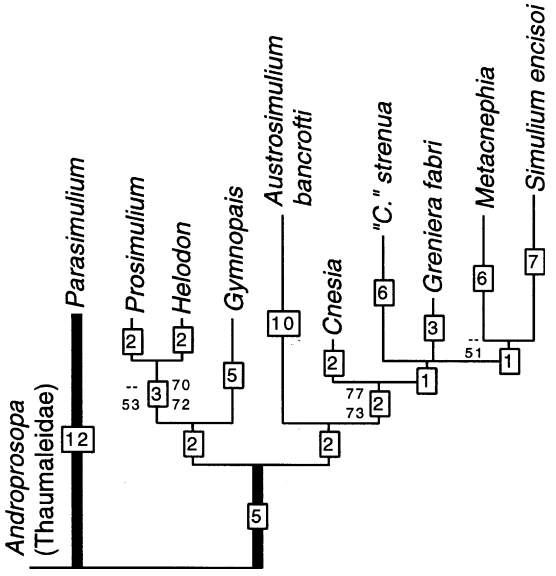


Fig. 3. Most parsimonious tree resulting from the application of a 5/20/1 (nt1/nt2/nt3) codon weighting scheme to PEPCK nucleotides (TL = 612, CI = 0.592, and RC = 0.257). Bootstrap support for clades appearing in 50% or more of the 1000 bootstrap replicates conducted using maximum parsimony and codon weighting bootstrap (averaged over different schemes) are shown above those using minimum evolution and neighbor-joining criteria. Branch lengths (first and second codon positions only) are shown directly on branches. Subfamilies and tribes depicted by thickened solid and hatched lines, respectively. Subfamily and tribal bootstrap support not shown.

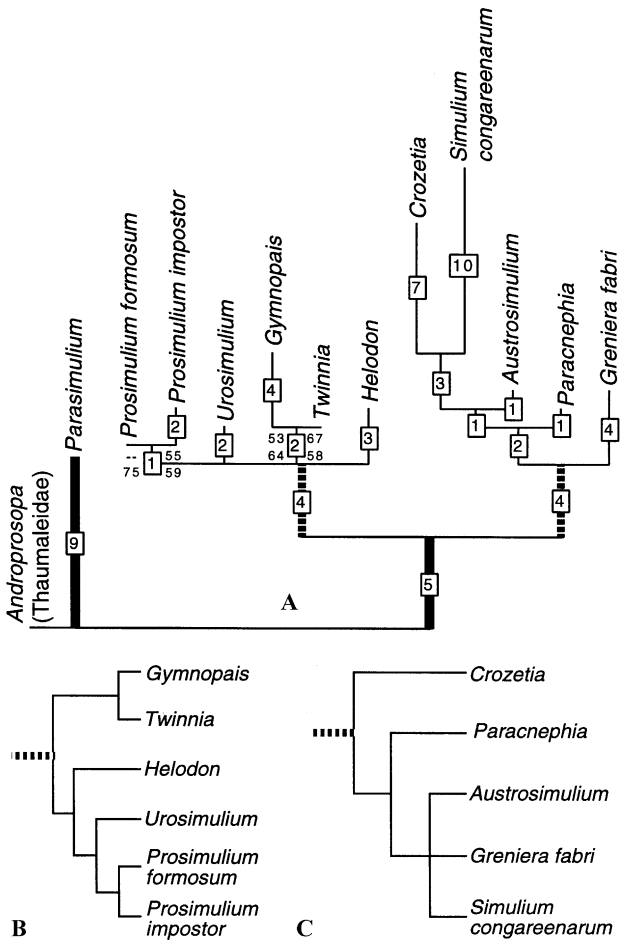


Fig. 4. Results from analysis of dopa decarboxylase sequences. (A) Consensus of two most parsimonious trees found using maximum parsimony criteria (TL = 663, CI = 0.575, and RC = 0.268). Maximum parsimony and codon weighting bootstrap (avg.) scores of 50 or greater are shown to the right of nodes and minimum evolution and neighbor joining bootstrap scores of 50 or greater are shown to the left of nodes. (B) Preferred relationships among Prosimuliini as inferred from distance analysis of nucleotides. (C) Consensus of 90 MPTS found from amino acid sequences (TL = 87; CI = 0.885, and RC = 0.721). *G. fabri* and *S. congareenarum* paired in 54% of the minimum evolution bootstrap replicates. Subfamily and tribal bootstrap support not shown.

*Paracnephia*/"*Cnephia*" *strenua*. Fig. 7 is a majority rule consensus of these three trees.

## 4. Discussion

Intratribal relationships within Prosimuliini and Simuliini remain largely unresolved. The loci examined in this study do not provide significantly greater phylogenetic resolution than does morphology. In a taxonomical system only slightly more conservative than the one most widely accepted today, e.g., Crosskey (1988), most intratribal groupings recovered by these molecular data could easily be considered genera. The following sections chronicle the major findings of this research, as deemed by degree of node support (concordance, bootstrap, and decay scores) among molecular-based trees and concordance with morphological data.

### *4.1. Previously recognized groups recovered with strong support*

Analyses of 28S and/or EF-1α provide strong bootstrap and/or decay support for several previously hypothesized intra- and intergeneric relationships. These groups and the genes that support them are as follows: *Gymnopais* + *Twinnia* (28S, 12S, and DDC), *Cnesia* + *Gigantodax* (EF-1α), *Stegopterna* + *Mayacnephia* (28S, 12S, and EF-1α), *Paracnephia* + *Procnephia* (28S, EF-1α), and "*Cnephia*" *aurantiacum* + "*Cnephia*" *strenua* + "*Cnephia*" *tonnoiri* (= "*Cnephia*" *aurantiacum* species group *sensu* Mackerras and Mackerras (1948)) (28S, EF-1α), and *Simulium s.l.* (28S, EF-1α).

A sister group relationship between *Gymnopais* and *Twinnia* has been inferred on numerous occasions using morphological evidence (Borkent and Wood, 1986; Currie, 1988; Py-Daniel, 1990; Rubtzov, 1956, 1974; Wood, 1978). Currie (1988) and Py-Daniel (1990) propose 8 and 15 characters, respectively, as synapomorphies of this lineage. Adults of *Cnesia* and *Gigantodax* share apomorphic features in the terminalia of both sexes (Wygodzinsky and Coscarón, 1973). *Stegopterna*

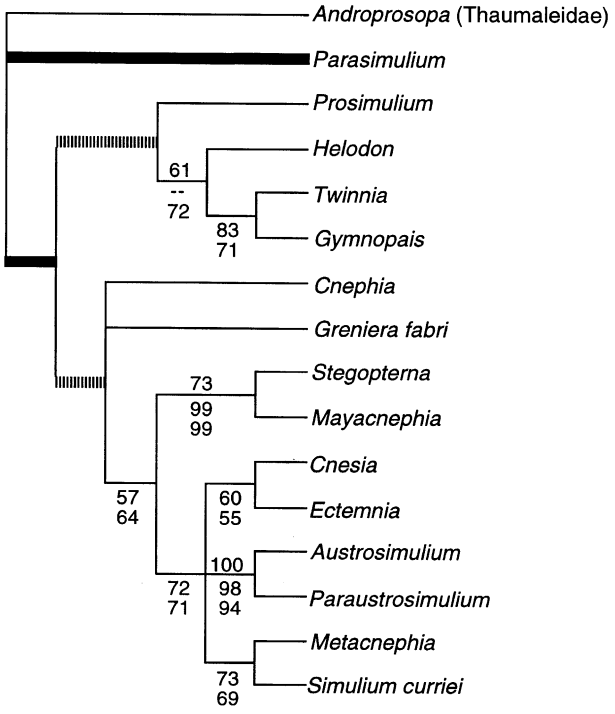


Fig. 5. Neighbor-joining tree inferred from elision matrix of 12S nucleotides. Maximum parsimony bootstrap scores of 50 or greater are shown above nodes and minimum evolution and neighbor-joining bootstrap scores of 50 or greater shown below nodes. Subfamily and tribal bootstrap support not shown.

and *Mayacnephia sensu* Currie (including *Tlalocomyia*) are the only genera in the family, other than *Parasimulium*, in which larvae (of most species) have a continuous transverse bulge (mid-ventral bulge, MVB) on the venter of abdominal segment VIII (Currie, 1988; Stone and Jamnback, 1955; Wygodzinsky and Coscarón, 1973; Wygodzinsky and Díaz Nájera, 1970). Rubtzov (1962) and Crosskey and Howard (1997) consider *Paracnephia* and *Procnephia* congeneric, but the morphological basis is unclear. Although Mackerras and Mackerras (1948) provided no synapomorphies for the "*Cnephia*" *aurantiacum* species group, several putative ones have since been identified (J.K. Moulton, unpublished). The recovery of a monophyletic *Simulium s.l.* is in contrast with results from inferences from partial sequences of 16S rDNA (Tang et al., 1996).

### 4.2. Previously unrecognized groups recovered with strong support

Bootstrap analyses of 28S and/or EF-1α also support monophyly of three groups of closely related species that are in the process of being formally described by the author as new genera. Two are endemic to Australia, making up part of the Australian "*Cnephia*" or Prosimuliini of authors. These are "*Cnephia*" y + "*Cnephia*" sp. nr. *terebrans* and "*Cnephia*" 'S. x. (East)' + "*Cnephia*" 'S. x. (West)'. Members of both of these groups are quite distinct morphologically and share several synapomorphies, especially in the preimaginal stages (J.K. Moulton, unpublished). The third group is comprised of three species currently assigned to three different genera: *P. anthracinum*, "*Cnephia*" *pilfreyi*, and "*Austrosimulium*" *colboi*. This clade is strongly supported by distance analyses of 28S sequences. These species share numerous features, such as the short second antennal article of the larva, a simple *Simulium*-like hypostoma partially shielded by the ventral hypostomal wall (Currie, 1986) (= laminar projection *sensu* Py-Daniel, 1990), short terminal pupal spines, dististyle with three or more apical teeth, and a geniculate, medially divided median sclerite (male terminalia) that serve as synapomorphies linking them with *Cnesiamima atroparva* (Wygodzinsky and Coscarón, 1973). The similar structure of the pupal gills of *C. atroparva*, "*A.*" *colboi*, and *P. anthracinum* is hypothesized to be a synapomorphy linking these three taxa. *P. anthracinum* and "*A.*" *colboi* should be considered sister species within *Paraustrosimulium*.

*Greniera denaria* is inferred as the sister group to *Stegopterna* + *Mayacnephia* in parsimony and minimum evolution analyses of 28S, distance analyses of EF-1α, and distance analyses of the 28S + EF-1α matrix. When concatenated 28S and EF-1α sequences are analyzed simultaneously, this trio had bootstrap scores of 64 and 68 in minimum evolution and neighbor-joining analyses (Fig. 6). These genera have superficial similarities to one another, especially in the larval stage, but convincing synapomorphies linking them are currently lacking.

All analyses of 28S, parsimony and NJ analyses of EF-1α, and codon weighting analyses of PEPCK in which second positions are given five or more times the weight of first positions support a sister group relationship between *Metacnephia* and *Simulium s.l.* This arrangement has not been previously hypothesized, but potential synapomorphies for it have been identified (D.C. Currie and P.H. Adler, personal communication).

Analyses of 28S provide strong support for a sister relationship between the *Simulium* (*Hellichiella*) *subexcisum* group, of which *S. congareenarum* is a member, to the remainder of *Simulium s.l.* The *Simulium annulum* group, of which *S. canonicolum* is a member, has consistently been placed within *Hellichiella* (Crosskey, 1988; Rubtzov, 1956). Analyses of 28S sequences, however, strongly suggest this group is more closely related to the other *Simulium* groups examined. It is noteworthy that the species of *Simulium* that Dunbar (1967) claimed as the closest to *Cnephia*, *Simulium anatinum*, belongs to the *S.* (*H.*) *subexcisum* group and is the sister species of *S. congareenarum*. Relationships among *Simulium* exclusive of *S. congareenarum* are not well resolved by these data.

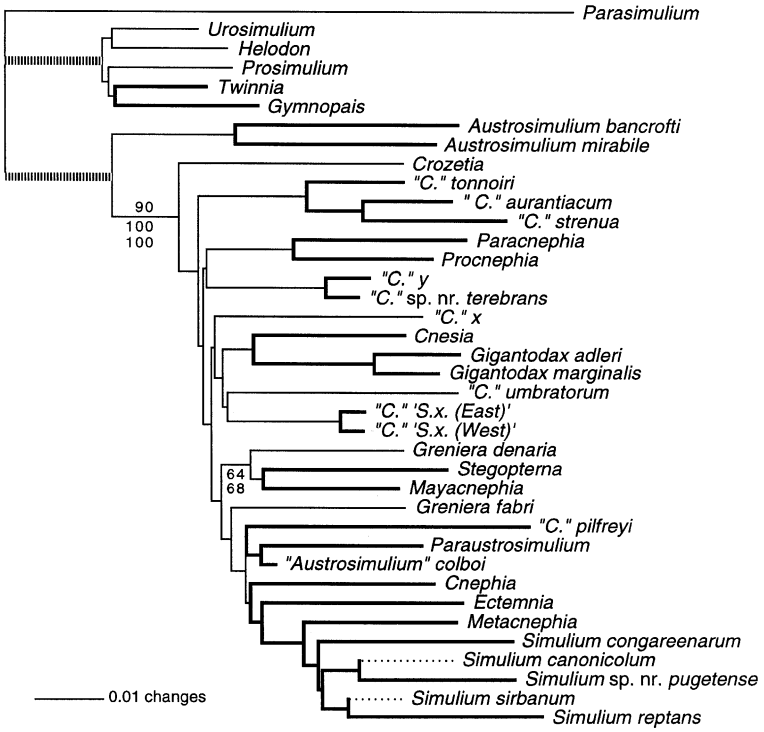


Fig. 6. Tentative "best estimate" of simuliid relationships based upon molecular data analyzed herein. This phylogram is the neighbor joining tree from simultaneous analysis of 28S and EF-1α nucleotides using HKY85-corrected distances with clades strongly supported by 28S constrained to be monophyletic (shown by thickened lines). The *Greniera denaria*/*Stegopterna*/*Mayacnephia* clade is the only intratribal one not constrained to be monophyletic that was recovered in 50% or more of the bootstrap replicates conducted during maximum parsimony, minimum evolution, or neighbor-joining searches. The position of the root in both tribes differs from that of some trees inferred from other genes examined herein and expectations based upon morphological evidence. Subfamily and tribal bootstrap support not shown. The distal outgroup, *A. gillespieae*, was removed following inference to facilitate visualization of short internal nodes within Simuliidae.

### *4.3. Previously recognized or hypothesized groups recovered with weak support*

A sister group relationship between *Prosimulium* + *Urosimulium* was recovered from distance analyses of 28S and DDC nucleotide sequences. This arrangement is corroborated by weak cytological evidence (Frizzi et al., 1970; Rothfels, 1979) and similarities in the larva and pupal gill (Bernard et al. (1972); Crosskey, 1969, 1981, 1988). These authors consider *Urosimulium* as the *Prosimulium aculeatum* species group, whereas other authors consider it a distinct genus (Contini, 1963, 1966; Currie, 1988; Rivosecchi, 1978; Rubtzov, 1974). Currie (1988) hypothesized that *Urosimulium* may be the sister group to a monophyletic group containing *Levitinia*, *Twinnia*, and *Gymnopais*. Until more compelling evidence is presented, *Urosimulium* would best be considered *incertae cedis* and as a distinct genus.

Trees from all analyses of 28S sequences, analyses of EF-1α in which "*Cnephia*" *pilfreyi* and *Paraustrosimulium* are constrained to be monophyletic support the monophyly of a clade containing *Cnephia*, *Ectemnia*, *Greniera fabri*, *Metacnephia*, and *Simulium s.l.* Bootstrap support from 28S neighbor-joining analysis is particularly strong, i.e., 97%. With the exception of *Greniera fabri*, this clade is conducive with previous authors' notions that *Simulium* is most recently derived from a *Cnephia*-like ancestor (Dumbleton, 1963, 1972; Shewell, 1958). Dunbar (1967) claims that the closest members of *Cnephia* and *Eusimulium*, *Cnephia dacotensis* and *Eusimulium anatinum* Wood, differ in only one inversion in the short arm of chromosome II. The importance of this chromosomal similarity, however, is unclear. Adults of *Cnephia* and *Ectemnia* superficially resemble one another, and at one time they were considered congeneric (Stone and Jamnback, 1955). All of these taxa, with the exception of *Greniera fabri* and *Simulium* (*Hellichiella*) (in part) have a larval hypostoma comprised of small uniformly sized teeth that are arranged more or less linearly. The importance of this character, however, is also questionable due to homoplasy; similar hypostomae occur in *Austrosimulium*, "*Cnephia*" X, "*Cnephia*" *pilfreyi* and allies, and *Lutzsimulium*.

### *4.4. Well-supported clades strongly discordant with morphological evidence*

Two clades recovered with significant support are strongly at odds with morphological evidence: (1) *Au-*

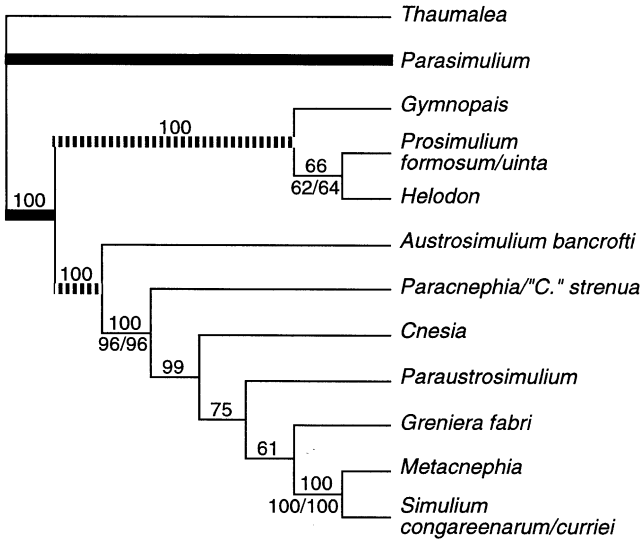


Fig. 7. Most parsimonious tree found from maximum and codon weighting (of PEPCK and DDC sequences) (TL = 3265, CI = 0.698, and RC = 0.342). Maximum parsimony bootstrap scores 50% and greater are shown above nodes and minimum evolution and neighbor-joining bootstrap scores 50% or greater shown below nodes. The minimum evolution and neighbor joining trees were identical and differed from the most parsimonious tree only in that "*Cnesia*" and "*Paracnephia*" appeared as sister taxa; this grouping exhibited bootstrap scores of 56 and 58, respectively, in these analyses. Subfamily and tribal bootstrap support not shown.

*strosimulium s. str.* as the sister group to the remainder of Simuliini and (2) "*Cnephia*" *pilfreyi* as the sister group to *Ectemnia*. The former arrangement is supported by neighbor-joining analysis of 28S sequences, parsimony and distance analyses of EF-1α nucleotides, and parsimony and distance analyses of PEPCK nucleotides. As one would expect, combining these data in a total evidence approach results in very strong support for this arrangement. In some cases, this topology is more strongly supported than is monophyly of the tribe.

Monophyly of Simuliini exclusive of *Austrosimulium* is not supported by any morphological data of which I am aware. Morphological data suggests that *Austrosimulium* is most closely related to *Paraustrosimulium* and allied species (see above) and perhaps having closer ties to *Simulium*. *Austrosimulium* shares numerous features with "*Cnephia*" *pilfreyi* and allies (see above). Hence, several authors have hypothesized these groups are closely related (Coscarón, 1985; Edwards, 1931; Wygodzinsky and Coscarón, 1962). The basal position of *Austrosimulium* may be attributable to long branch attraction (Felsenstein, 1978), as the *Austrosimulium* branch and the one connecting Prosimuliini and Simuliini are among the three longest within 28S trees. An attempt to shorten the *Austrosimulium* branch proved unsuccessful in spite of including representatives from both putative subgenera. Thus, the small degree of non-synonymous variation observed within PEPCK among simuliids was evidently no guarantee that the few variable amino acid positions were not multiply substituted.

Strong support for a sister group relationship between "*Cnephia*" *pilfreyi* and *Ectemnia* in analyses of EF-1α sequences contradicted the 28S (ME/NJ analyses only) placement of "*Cnephia*" *pilfreyi* within a group containing *Paraustrosimulium* and "*Austrosimulium*" *colboi*. Although bootstrap values are greater for the EF-1α arrangement, the 28S topology is preferred because it is corroborated by convincing morphological evidence (see above). Analyses of EF-1α and other genes, such as COI, in which synonymous variation comprises virtually all observed variation, are open to problems stemming from nucleotide biases.

### *4.5. Preferred estimate of intratribal relationships within Simuliidae*

The minimum evolution tree (Fig. 6) recovered from combined analysis of 28S and EF-1α nucleotides with clades strongly supported by 28S constrained to be monophyletic is presented as the most reasonable estimate of simuliid relationships. This tree was chosen as such because it is concordant with and much more densely sampled than the total evidence tree and it contains several features that are at least partially concordant with traditional views based upon morphology.

### *4.6. Phylogenetic utility and the rapid divergence problem*

The genes examined in this study, plus several (18S, COI, and Wint-1) that were removed from consideration during the early phases of this study, represent the bulk of our current molecular systematics arsenal. Although these genes represent both major genomes and encode vastly different products, all are suboptimal indicators of most intratribal relationships within Simuliidae. Of the three markers dismissed early in this study, 18S, wingless (Wint-1), and COI, 18S was nearly invariant, wingless was too conserved with respect to non-synonymous variation, and COI was virtually invariant at non-synonymous sites and positively misleading due to saturation of silent sites (J.K. Moulton, unpublished). The latter assumption was based upon high levels of divergence at third positions and the failure to recover a monophyletic Simuliidae or its constituent subfamilies and tribes.

To a lesser extent, these problems were also observed within the genes examined herein. Variation in 28S [pair-wise divergence = 2–6% among Simuliinae] and non-synonymous variation in DDC [pair-wise divergence = 1–10% among Simuliinae] and PEPCK [pair-wise divergence = 1–7% among Simuliinae] was insufficient to supply an adequate number of informative sites within simuliids, particularly at levels lower than tribe, and synonymous variation within EF-1α, DDC, and PEPCK was at or near saturation. Greater similarity

between the outgroup and derived simuliid taxa in pairwise comparisons of EF-1α, DDC, and PEPCK third positions attests to the high level of saturation and/or codon biases present in this subset of characters. It is remarkable, however, that EF-1α was able to recover subfamilial and tribal relationships within Simuliidae with as much or even greater bootstrap support than did any of the other genes examined herein (Moulton, 2000).

Divergence within 12S among the black fly taxa examined ranged from 4 to 22%, with most intratribal divergences ranging from 6 to 10%. This amount is 2–3 times greater than that observed in similar comparisons of 28S, DDC (nt1 + 2), and PEPCK (nt1 + 2). Trees inferred from 12S have considerably greater node support on average than do ones inferred from the other two genes. The weakness of 12S, however, lies with its small size, difficult alignment, and strong adenine/thymine bias. Adenine and thymine comprise approximately 78% of the nucleotides in this region of 12S within Simuliidae (Moulton, 2000) and transversions between them account for 44% (134 of the 305 unambiguous changes) of the observed variation. These adenine/thymine transversions are subject to multiple hits (Simon et al., 1994).

The inability of 18S, PEPCK, DDC, EF-1α, and 12S, particularly DDC and PEPCK, to adequately resolve most simuliid relationships is troubling. Dopa decarboxylase and PEPCK are considered to be two of the most promising indicators of Mesozoic-age divergences based upon their ability to reconstruct known phylogenies of lepidopteran groups that evolved during that time (Fang et al., 1997; Friedlander et al., 1996). Fossils of extant genera within Simuliini appear during the latter half of this period (D.C. Currie, personal communication; Jell and Duncan, 1986). The suboptimal phylogenetic utility observed for DDC and PEPCK within Simuliidae is perhaps due to rapid divergence events within simuliid tribes and the relatively small size and mode and tempo of evolution of these genes. These genes, or at least the regions of them examined to date, seem incapable of accumulating meaningful, e.g., nonsynonymous, mutations quickly enough to track rapid divergence events.

Inference of relationships within ancient complex clades requires the use of considerable amounts of sequence data (Hillis, 1996). This is particularly true if the group experienced one or more bouts of rapid cladogenesis. Such an endeavor also requires the use of one or more genes containing characters that evolved at different rates and form natural partitions having at least some complimentarity, if not outright concordance. Individual nuclear coding genes capable of supplying multiple complimentary data partitions appear to be absent from the current repertoire of insect molecular systematics markers. Moreover, if such genes were available, it is unlikely that fragments from them large enough to span several functional domains would be utilized.

As a means of testing these hypotheses, I am currently evaluating the utility of the multifunctional genes CAD and GART for resolving difficult relationships within Simuliidae. Within flies, these genes are large, single copy, nearly uninterrupted (few and generally small introns) stretches of coding sequence encoding multiple domains containing conserved areas useful for primer design. Complete sequences of both genes are known from *Drosophila* (Adams et al., 2000; Freund and Jarry, 1987; Henikoff and Eghtedarzadeh, 1987) and GART is known from *Chironomus* (Clark and Henikoff, 1992). Results from analyses of preliminary data consisting of a 2-kb portion of GART and a 1-kb portion of CAD from several simuliid taxa are encouraging. Of particular interest is a more reasonable placement of *Austrosimulium* nearer morphologically similar genera. Hence, better phylogenetic resolution within Simuliidae and other rapidly diverged taxa awaits the discovery and implementation of better-suited markers.

## Acknowledgments

D.R. Maddison provided logistical and technical support for this research. D.R. Maddison, P.H. Adler, D.G. Mead, R. DeSalle, and an anonymous reviewer provided excellent reviews of earlier drafts of this manuscript that improved it significantly. J. Prince (University of Western Australia) and P. Cranston (UC-Davis) were gracious hosts whilst I was in Australia collecting simuliids. The following persons supplied material used in this study: P.H. Adler (Clemson University), D.B. Bedo (DPI-Queensland), P. Cranston (CSIRO-Canberra), D.C. Currie (University of Toronto), D.M. Wood (ECORC-retired), R.W. Crosskey (NHM-retired), S. Coscarón (University of La Plata), R. Miranda Esquivel (University of La Plata), F.C. de Moor (Albany Museum, Grahamstown, South Africa), and B.E. Marshall (University of Zimbabwe). This research was supported by NSF-DDIG (DEB-9520613) awarded to J.K.M., a Flinn Foundation grant awarded to J.K.M., NSF Research Training Group (RTG) Grant in the Analysis of Biological Diversification (DIR-9113362) awarded to the University of Arizona, and NSF (DEB-9420219) awarded to D.R. Maddison.

## References

Abouheif, E., Zardoya, R., Meyer, A., 1998. Limitations of metazoan 18S rRNA sequence data: implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. J. Mol. Evol. 47, 394–405.

Adams, M.D. et al., 2000. The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195.

Bernard, M.R., Grenier, P., Bailly-Choumara, H., 1972. Description de *Prosimulium* (*Prosimulium*) *faurei* n. sp. (Diptera: Simuliidae).

Cahiers O.R.S.T.R.O.M., séries Entomologie Médicale et Parasitologie 10, 63–68.
Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1998. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. Mol. Biol. Evol. 15, 1719–1727.
Borkent, A., Wood, D.M., 1986. The first and second larval instars and the egg of *Parasimulium stonei* Peterson. Proc. Entomol. Soc. Wash. 88, 287–296.
Bremer, K., 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42, 795–803.
Brower, A.V.Z., DeSalle, R., 1998. Patterns of mitochondrial and nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. Insect Mol. Biol. 7, 73–82.
Caterino, M.S., Cho, S., Sperling, F.A.H., 2000. The current state of insect molecular systematics: a thriving Tower of Babel. Ann. Rev. Entomol. 45, 1–54.
Cho, S., Mitchell, A., Regier, J.C., Mitter, C., Poole, R.W., Friedlander, T.P., Zhao, S., 1995. A highly conserved nuclear gene for low-level phylogenetics: *elongation factor*-1α recovers morphology-based tree for heliothine moths. Mol. Biol. Evol. 12, 650–656.
Clark, D.V., Henikoff, S., 1992. Unusual organizational features of the *Drosophila*-Gart locus are not conserved within Diptera. J.. Mol. Evol. 35, 51–59.
Contini, C., 1963. Nuovi reperti e note ecologiche di alcuni simulidi della Sardegna e descrizione di *Urosimulium stefanii* n. gen., n. sp. Memorie della Societa Entomologica Italiana 14, 87–97.
Contini, C., 1966. *Urosimulium juccii* n. sp. (Diptera: Simuliidae), nuova specie de Simulide della Sardegna. Riv. Parasitol. 27, 269–291.
Coscarón, S., 1985. On *Cnesiamima atroparva* Edwards larva and its significance to interpret the phylogenetic relationships of notogeic black flies (Insecta, Diptera, Simuliidae). Bol. Soc. Biol. Concep. 56, 67–70.
Crosskey, R.W., 1969. A re-classification of the Simuliidae (Diptera) of Africa and its islands. Bull. Brit. Mus. Nat. Hist. (Entomology) (Supplement 14), 195 pp.
Crosskey, R.W., 1981. Simuliid taxonomy—the contemporary scene. In: Laird, (Ed.), Blackflies: The Future for Biological Methods in Integrated Control. Academic Press, New York, pp. 3–18, 399 pp.
Crosskey, R.W., 1988. An annotated checklist of the world's black flies (Diptera: Simuliidae). In: Kim, Merritt, (Eds.), Black Flies: Ecology, Population Management and Annotated World List. Penn. State Press, pp. 425–520.
xv + 528 pp.
Crosskey, R.W., Howard, T.M., 1997. A New Taxonomic and Geographical Inventory of World Blackflies (Diptera: Simuliidae). Natural History Museum, London.
Currie, D.C., 1986. An annotated list of and keys to the immature black flies of Alberta (Diptera: Simuliidae). Mem. Entomol. Soc. Can., 134, 1–90 pp.
Currie, D.C., 1988. Phylogeny of primitive Simuliidae (Insecta: Diptera: Culicomorpha). Ph.D. dissertation, University of Alberta, Edmonton, Canada.
Dumbleton, L.J., 1963. The classification and distribution of the Simuliidae (Diptera) with particular reference to the genus *Austrosimulium*. NZ J. Sci. 6, 320–357.
Dumbleton, L.J., 1972. The genus *Austrosimulium* Tonnoir (Diptera: Simuliidae) with particular reference to the New Zealand fauna. NZ J. Sci. 15, 480–584.
Dunbar, R.W., 1967. The salivary gland chromosomes of six closely related black flies near *Eusimulium congareenarum* (Diptera: Simuliidae). Can. J. Zool. 45, 377–396.
Edwards, F.W., 1931. Simuliidae. In: Diptera of Patagonia and South Chile, Part 2, pp. 121–154.
Fang, Q.Q., Cho, S., Regier, J.C., Mitter, C., Matthews, M., Poole, R.W., Friedlander, T.P., Zhao, S., 1997. A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). Syst. Biol. 46, 269–283.
Felsenstein, J., 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.
Foster, P.G., Jermiin, L.S., Hickey, D.A., 1997. Nucleotide composition bias affects amino acid content in proteins coded by animal mitochondria. J. Mol. Evol. 44, 282–288.
Freund, J.N., Jarry, B.P., 1987. The rudimentary gene of *Drosophila melanogaster* encodes four enzymatic functions. J. Mol. Biol. 193, 1–13.
Friedlander, T.P., Regier, J.C., Mitter, C., 1992. Nuclear gene sequences for higher-level phylogenetic analysis: 14 promising candidates. Syst. Biol. 41, 483–489.
Friedlander, T.P., Regier, J.C., Mitter, C., 1994. Phylogenetic information content of five nuclear gene sequences in animals: initial assessment of character sets from concordance and divergence studies. Syst. Biol. 43, 511–525.
Friedlander, T.P., Regier, J.C., Mitter, C., Wagner, D.L., 1996. A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinase tracks Mesozoic-age divergences within Lepidoptera (Insects). Mol. Biol. Evol. 13, 594–604.
Frizzi, G., Lecis, A., Contini, C., 1970. Sex determination in Urosimulium stefanii Cont. (Diptera: Simuliidae). Zeitschrift fur Zoologische Systematic und Evolutionforschung 8, 154–159.
Gilbert, D., 1992. SeqApp, version 1.9a157 (available thru anonymous ftp to ftp.bio.indiana.edu, in folder /molbio/seqapp, as seqapp.hqx).
Graybeal, A., 1994. Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. Syst. Biol. 43, 174–193.
Halanych, K.M., 1998. Considerations for reconstructing metazoan history: signal, resolution, and hypothesis testing. Am. Zool. 38, 929–941.
Halanych, K.M., Robinson, T.J., 1999. Multiple substitutions affect the phylogenetic utility of cytochrome *b* and 12S rDNA data: examining a rapid radiation in Leporid (Lagomorpha) evolution. J. Mol. Evol. 48, 369–379.
Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21, 160–174.
Hasegawa, M., Hashimoto, T., 1993. Ribosomal RNA trees misleading? Nature 361, 23.
Henikoff, S., Eghtedarzadeh, M.K., 1987. Conserved arrangement of nested genes at the *Drosophila* GART locus. Genetics 117, 711–726.
Hickson, R., 1993. Sequences and signals: evolutionary histories of New Zealand skinks. Ph.D. Thesis, Massey University, Palmerston North, New Zealand.
Hillis, D.M., 1996. Inferring complex phylogenies. Nature 383, 130–131.
Hwang, U.W., Kim, W., Tautz, D., Friedrich, M., 1998. Molecular phylogenetics at the Felsenstein zone: approaching the Strepsiptera problem using 5.8S and 28S rDNA sequences. Mol. Phylogenet. Evol. 9, 470–480.
Irwin, D.M., Kocher, T.D., Wilson, A.C., 1991. Evolution of the cytochrome-*b* gene of mammals. J. Mol. Evol. 32, 128–144.
Jell, P.A., Duncan, P.M., 1986. Invertebrates, mainly insects, from the freshwater, Lower Cretaceous, Koonwarra Fossil Bed (Korumburra Group), South Gippsland,Victoria. Mem. Assoc. Austral. Paleontol. 3, 111–205.
Kalugina, N.S., 1991. New Mesozoic Simuliidae and Leptoconopidae and the origin of blood- sucking in lower dipterans. Paleontologicheskii Zhurnal 1, 69–80.
Kalugina, N.S., Kovelev, V.G., 1985. Jurrasic Diptera of Siberia. Nauka, Moscow, 198 pp.
Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial

DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86, 6196–6200.
Kraus, F., Miyamoto, M.M., 1991. Rapid cladogenesis among the pecoran ruminants: evidence from mitochondrial DNA sequences. Syst. Zool. 40, 117–130.
Ledje, C., Arnason, U., 1996. Phylogenetic relationships within caniform carnivores based upon analyses of the mitochondrial 12S rRNA gene. J. Mol. Evol. 43, 641–649.
Li, W.H., Graur, D., 1991. Fundamentals of Molecular Evolution. Sinauer Associates, Sunderland, MA.
Liu, H., Beckenbach, A.T., 1992. Evolution of the mitochondrial cytochrome oxidase II gene among ten orders of insects. Mol. Phylogenet. Evol. 1, 41–52.
Mackerras, I.M., Mackerras, M.J., 1949. Revisional notes on Australasian Simuliidae (Diptera). Proc. Linn. Soc. N.S.W. 73, 372–405.
Mackerras, I.M., Mackerras, M.J., 1952. Notes on Australasian Simuliidae (Diptera). III. Proc. Linn. Soc. N.S.W. 77, 104–113.
Mackerras, M.J., Mackerras, I.M., 1948. Simuliidae (Diptera) from Queensland. Austral. J. Sci. Res. (B) 1, 231–270.
Mackerras, M.J., Mackerras, I.M., 1950. Notes on Australasian Simuliidae (Diptera). II. Proc. Linn. Soc. N.S.W. 75, 167–187.
Mackerras, M.J., Mackerras, I.M., 1955. Notes on Australasian Simuliidae (Diptera). IV. Proc. Linn. Soc. N.S.W. 80, 105–112.
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, Version 3.0. Sinauer Associates, 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.
Mitchell, A., Cho, S., Regier, J.C., Mitter, C., Poole, R.W., Matthews, M., 1997. Phylogenetic utility of *elongation factor*-1α in Noctuoidea (Insecta: Lepidoptera): the limits of synonymous variation. Mol. Biol. Evol. 14, 381–390.
Moulton, J.K., 1997. Molecular systematics of the Simuliidae (Diptera: Culicomorpha). Ph.D. Thesis. University of Arizona, Tucson, AZ.
Moulton, J.K., 2000. Molecular sequence data resolves basal divergences within Simuliidae (Diptera). Syst. Entomol. 25, 95–113.
Nei, M., 1978. Molecular Evolutionary Genetics. Columbia University Press, New York.
Norris, D.E., Klompen, J.S.H., Black, W.C., 1999. Comparison of the mitochondrial 12S and 16S ribosomal DNA genes in resolving phylogenetic relationships among hard ticks (Acari: Ixodidae). Ann. Entomol. Soc. Am. 92, 117–129.
Orti, O., Meyer, A., 1996. Molecular evolution of ependymin and the phylogenetic resolution of early divergences among euteleost fishes. Mol. Biol. Evol. 13, 556–573.
Py-Daniel, V., 1990. Revisao da Tribo Prosimuliini seg. Crosskey (Diptera, Culicomorpha) nas regioes Zoogeographicas Neartica E Neotropical (larvas and pupas). Ph.D. Thesis, Instituto Nacional de Pesquisas da Amazonia, Manaus, Brazil.
Regier, J.C., Schultz, J.W., 1998. Molecular phylogeny of arthropods and the significance of the Cambrian "explosion" for molecular systematics. Am. Zool. 38, 918–928.
Rivosecchi, L., 1978. Simuliidae: Diptera, Nematocera. Fauna d'Italia 13, 535 pp.
Roger, A.J., Sandblom, O., Doolittle, W.F., Philippe, H., 1999. An evaluation of elongation factor-1α as a phylogenetic marker for eukaryotes. Mol. Biol. Evol. 16, 218–233.
Rothfels, K., 1979. Cytotaxonomy of black flies (Simuliidae). Ann. Rev. Entomol. 24, 507–539.
Rubtzov, I.A., 1956. *Fauna of the USSR.* (Insecta, Diptera). Black flies (family Simuliidae). Zoological Institute of the Academy of Science USSR (n.s. 64), 6, 1–860 (translated from Russian).
Rubtzov, I.A., 1962. Simuliid genera in the fauna of the Ethiopian Region. Zool. Zh. 41, 1488–1502 (in Russian, with English summary).
Rubtzov, I.A., 1974. Evolution, phylogeny, and classification of the family Simuliidae (Diptera). British Library Lending Division, Wetherby. 83 pp (translated from Russian).
Shewell, G.E., 1958. Classification and distribution of arctic and subarctic Simuliidae. In: Proceedings of the 10th International Congress on Entomology, Montreal, 1956, vol. 1, pp. 635–643.
Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.
Stone, A., Jamnback, H.A., 1955. The black flies of New York State. NY State Mus. Bull., 349, 144 pp.
Swofford, D.L., 1996. PAUP*. Version 4.0d52 (a beta test version, used with permission of D.L. Swofford; completed version 4.0 to be distributed by Sinauer Associates, Sunderland, MA).
Swofford, D.L., 1997. PAUP*. Version 4.0d54 (a beta test version, used with permission of D.L. Swofford; completed version 4.0 to be distributed by Sinauer Associates, Sunderland, MA).
Tang, J., Toe, L., Pruess, K., Cupp, E.W., Unnasch, T.R., 1996. Molecular phylogeny and typing of blackflies (Diptera: Simuliidae) that serve as vectors of human or bovine onchocerciasis. Med. Vet. Entomol. 10, 228–234.
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 Research 24, 4876–4882.
Vossbrinck, C.R., 1989. A 28S ribosomal RNA phylogeny of certain cyclorrhaphous Diptera based upon a hypervariable region. Syst. Entomol. 14, 417–431.
Wheeler, W.C., Gatesy, J., DeSalle, R., 1995. Elision: a method for accommodating mulitple molecular sequence alignments with alignment-ambiguous sites. Mol. Phylog. Evol. 4, 1–9.
Wood, D.M., 1978. Taxonomy of the Nearctic species of *Twinnia* and *Gymnopais* (Diptera: Simuliidae) and a discussion of the ancestry of the Simuliidae. Can. Entomol. 110, 1297–1337.
Wygodzinsky, P., Coscarón, S., 1962. On the relationships and zoographical significance of *Austrosimulium anthracinum* (Bigot), a blackfly from southern South America. Pac. Insects 4, 235–244.
Wygodzinsky, P., Coscarón, S., 1973. A review of the Mesoamerican and South America black flies of the tribe Prosimuliini (Simuliinae, Simuliidae). Bull. Am. Mus. Nat. Hist. 151, 129–200.
Wygodzinsky, P., Díaz Nájera, A., 1970. Un nuevo genero de Simulido de la Republica Mexicana (Diptera: Simuliidae). Rev. Inv. Salud Publ. 30, 83–110.