

**UNIVERSIDADE DE SÃO PAULO  
FFCLRP - DEPARTAMENTO DE BIOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENTOMOLOGIA**

**Análises filogenéticas de Diptera (Insecta) com dados  
moleculares: genes, alinhamentos, amostragem e  
algoritmos.**

Ana Claudia Kaminski

Tese apresentada à Faculdade de  
Filosofia, Ciências e Letras de  
Ribeirão Preto - USP, como parte  
das exigências para obtenção do  
título de Doutora em Ciências -  
Área: Entomologia.

Ribeirão Preto / SP  
2006

**UNIVERSIDADE DE SÃO PAULO  
FFCLRP - DEPARTAMENTO DE BIOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENTOMOLOGIA**

**Análises filogenéticas de Diptera (Insecta) com dados  
moleculares: genes, alinhamentos, amostragem e  
algoritmos.**

Ana Claudia Kaminski

Orientador: Prof. Dr. Dalton de Souza Amorim

Tese apresentada à Faculdade de  
Filosofia, Ciências e Letras de  
Ribeirão Preto - USP, como parte  
das exigências para obtenção do  
título de Doutora em Ciências -  
Área: Entomologia.

Ribeirão Preto / SP  
2006

*Aos meus pais Carlos e Rosi, e minha irmã  
Andresa, por todo incentivo, amor, dedicação e  
confiança.*

## AGRADECIMENTOS

Àquele que conhece minha vida, minhas coisas boas e más, minhas dores, meus momentos de alegria, meus desapontamentos e esforços;

Ao Dr. Dalton de Souza Amorim por sua orientação, conhecimento, incentivo, dedicação, paciência e voto de confiança;

Ao Prof. Dr. Fernando Marques pelo auxílio nas análises e pelas valiosas discussões;

À grande amiga Maria Isabel Protti de Andrade Balbi, que em todos os momentos me acolheu com palavras amigas, sempre me suportando ao seu lado sem perder o carinho e seu “jeitinho iluminado”;

Aos companheiros de laboratório Charles Morphy, Adolfo Calor, Guilherme Ribeiro, Maria Virgínia Urso-Guimarães, Sarah Oliveira, Pâmela Adorno, Rafaela “Zirigui” Falaschi, Emanuelle Corbi Corrêa e Vera Lúcia Orlandini pela amizade, troca de idéias, sugestões e risadas;

À Priscila, Nelsinho, Fátima e Nelson Okano (*in memoriam*) e Juliana Martins, pessoas que têm muita importância na minha vida, a quem posso chamar de “segunda família”;

Aos amigos Marco Aurélio França, Nathália M. Monéa, Max C. Langer, Carla Cresoni e Thamara Leal, pessoas queridas que estiveram presentes em vários momentos importantes;

À Renata Andrade, pelo auxílio na preparação de documentos, e pelos chás de camomila nas horas de aperto;

Aos meus filhotes Bernardo, Bianca e Puppy, pela companhia em todos os momentos de alegria e tristeza, nos quais me deram apoio incondicional;

Ao Dr. João M. F. Camargo e à Dra. Zilá L. P. Simões pela orientação no início do curso;

Ao Departamento de Biologia da FFCL-RP, à Coordenação de Pós-Graduação e ao Conselho do Curso de Entomologia, em especial ao Dr. Carlos Garófalo, pelo apoio dispensado durante todas as atividades do curso;

Ao CNPq pela concessão da bolsa de estudos;

À quem, direta ou indiretamente, colaborou para a execução deste trabalho.

## SUMÁRIO

	Pág.
LISTA DE FIGURAS .....	v
LISTA DE TABELAS .....	viii
RESUMO .....	x
ABSTRACT .....	xi
1. INTRODUÇÃO .....	01
1.1. Diversidade e Sistemática de Diptera .....	01
1.2. Infraordens de Diptera .....	02
1.2.1. Tipulomorpha .....	02
1.2.2. Psychodomorpha .....	03
1.2.3. Bibionomorpha .....	03
1.2.4. Culicomorpha .....	04
1.2.5. Axymyiomorpha .....	04
1.2.6. Ptychopteromorpha .....	05
1.2.7. Blephariceromorpha .....	05
1.2.8. Brachycera .....	05
1.3. Estudos moleculares em Diptera .....	07
1.4. Análise de dados moleculares .....	08
2. OBJETIVOS .....	14
3. MATERIAL E MÉTODOS .....	15
3.1. Fontes das seqüências .....	15
3.2. Método geral de alinhamento .....	15
3.3. O gene nuclear 28S .....	17
4. RESULTADOS E DISCUSSÃO .....	20
4.1. Gene mitocondrial ribossomal 12S .....	20
4.1.1. Primeiro grupo de terminais para o gene mitocondrial ribossomal 12S .....	20
4.1.2. Segundo grupo de terminais para o gene mitocondrial ribossomal 12S .....	23
4.2. Gene mitocondrial ribossomal 16S .....	25

4.2.1. Primeiro grupo de terminais para o gene mitocondrial ribossomal 16S.....	25
4.2.2. Segundo grupo de terminais para o gene mitocondrial ribossomal 16S.....	27
4.3. Gene mitocondrial citocromo oxidase subunidade I (COI).....	30
4.3.1. Primeiro grupo de terminais para o gene mitocondrial COI.....	30
4.3.2. Segundo grupo de terminais para o gene mitocondrial COI.....	32
4.4. Gene mitocondrial citocromo oxidase subunidade II (COII).....	34
4.5. Robustez das topologias com dados dos genes mitocondriais 12S, 16S, COI e COII.....	36
4.6. Análise dos genes 18S e 28S.....	36
4.6.1. Análise do gene 18S.....	37
4.6.2. Análise do gene 28S.....	38
4.6.3. Análise combinada dos genes 18S e 28S.....	39
4.6.4. Análise combinada dos genes 18S e 28S com o aumento no número de táxons terminais, com inclusão de <i>missing data</i> .....	40
4.7. Análise do gene 28S com a inclusão da ordem Strepsiptera.....	41
4.8. Robustez das topologias com dados dos genes nucleares 18S e 28S.....	43
4.9. Diferentes métodos de alinhamento.....	45
4.10. Efeitos da amostragem de táxons sobre as topologias.....	50
4.10.1. Retirada de táxons terminais do grupo externo.....	50
4.10.2. Retirada de táxons terminais do grupo interno.....	53
4.11. Efeito da amostragem de diferentes regiões do gene sobre as topologias.....	57
4.11.1. Análise da região D7 do gene 28S.....	57
4.11.2. Análise da região D1 do gene 28S.....	60
4.11.3. Análise combinada das regiões D1 e D7 do gene 28S.....	64
4.12. A influência dos diferentes programas de alinhamento.....	67
4.13. Diferentes algoritmos de análise.....	69
5. CONSIDERAÇÕES FINAIS.....	71
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	73
ANEXOS.....	85

## LISTA DE FIGURAS

	Pág.
1. Fluxograma das análises realizadas com os três diferentes programas de alinhamento.....	18
2. Fluxograma representando a análise combinada das seqüências das regiões D1 e D7 do gene 28S, utilizando diferentes programas de alinhamento.....	19
3. Cladograma resultante da análise do primeiro grupo de terminais para o gene 12S, sob o parâmetro 111. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	22
4. Cladograma resultante da análise do segundo grupo de terminais para o gene 12S, sob o parâmetro 412. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	24
5. Cladograma resultante da análise do primeiro grupo de terminais para o gene 16S, sob o parâmetro 111. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	27
6. Cladograma resultante da análise do segundo grupo de terminais para o gene 16S, sob o parâmetro 111. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	29
7. Cladograma resultante da análise do primeiro grupo de terminais para o gene COI, sob o parâmetro de custos 4:2. Em cinza, os terminais que não pertencem a Diptera.....	31
8. Cladograma resultante da análise do segundo grupo de terminais para o gene COI, sob o parâmetro de custos 2:4. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	33
9. Cladograma resultante da análise do gene COII, sob o parâmetro de custos 2:4. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.....	35

10.	Cladograma resultante da análise do gene 18S, sob o parâmetro de custos 111.	38
11.	Árvore de consenso gerada pela análise do gene 28S, sob o parâmetro de custos 111.	39
12.	Cladograma resultante da análise combinada dos genes 18S e 28S, sob o parâmetro de custos 111.	39
13.	Cladograma resultante da análise combinada dos genes 18S e 28S, sob o parâmetro de custos 111, com a inclusão de <i>missing data</i> .	41
14.	Cladograma resultante da análise do gene 28S, sob o parâmetro de custos 111, com a inclusão de Strepsiptera no grupo externo. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.	43
15.	Trecho da seqüência da região D1 do gene 28S alinhado pelos programas MALIGN (1) e POY (2). Os caracteres autapomórficos estão marcados em azul, enquanto os caracteres que apresentam os cinco estados de caráter, em vermelho.	47
16.	Árvores encontradas para as matrizes geradas pelos programas CLUSTAL (A), MALIGN (B) e POY (C), com todos os táxons terminais.	49
17.	Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa com a retirada de sete dos dez táxons terminais do grupo externo (B).	51
18.	Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa com a retirada de sete dos dez táxons terminais do grupo externo (B).	52
19.	Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo externo (B).	53
20.	Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).	54
21.	Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).	55
22.	Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).	56
23.	Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa para a região D7 do gene 28S (B).	58



24.	Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e as encontradas pelo mesmo programa para a região D7 do gene 28S (B, C e D).	59
25.	Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa para a região D7 do gene 28S (B).....	60
26.	Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa para a região D1 do gene 28S (B).....	61
27.	Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa para a região D1 do gene 28S (B).....	62
28.	Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e as encontradas pelo mesmo programa para a região D1 do gene 28S (B, C e D).	63
29.	Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B).....	64
30.	Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e as encontradas em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B, C e D).....	66
31.	Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B).....	67
32.	Conjunto de topologias que representam as diferentes hipóteses de relações filogenéticas entre os grupos de dípteros basais encontradas no presente estudo.	68
33.	Cladogramas gerados pelos diferentes métodos de análise: (A) Neighbor-joining; (B) Maximum likelihood; e (C) Máxima Parcimônia, para a região D7 do gene 28S.....	70

## LISTA DE TABELAS

	Pág.
1. Número de pares de bases das seqüências utilizadas nas análises, antes dos alinhamentos.....	15
2. Custos para <i>indels</i> , transversões e transições utilizados nas análises dos genes 12S, 16S, 18S e 28S.....	16
3. Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o primeiro grupo de terminais para o gene 12S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.....	20
4. Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o segundo grupo de terminais para o gene 12S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.....	23
5. Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o primeiro grupo de terminais para o gene 16S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.....	26
6. Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o segundo grupo de terminais para o gene 16S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.....	28
7. Comprimento dos cladogramas obtidos para o primeiro grupo de terminais do gene COI, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.....	30
8. Comprimento dos cladogramas obtidos para o segundo grupo de terminais do gene COI, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.....	32
9. Comprimento dos cladogramas obtidos para o gene COII, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.....	34
10. Comprimento dos cladogramas obtidos para os genes 18S e 28S, separadamente e em análise combinada, sob os diferentes parâmetros de custos. Os números entre parênteses indicam o número de cladogramas igualmente parcimoniosos encontrados. Negrito indica o menor valor da ILD para o conjunto de dados.	37

11. Dados obtidos nas análises dos dados do gene 28S. A - análise com todos os táxons terminais; B - análise excluindo somente táxons do grupo externo; C - análise excluindo somente táxons do grupo interno; D - análise da região D1 do gene 28S, excluindo táxons dos grupos externo e interno; E - análise da região D7 do gene 28S, excluindo táxons dos grupos externo e interno; F - análise combinada das regiões D1 e D7 do gene 28S. Os números 1, 2 e 3 representam, respectivamente, os programas CLUSTAL, MALIGN e POY..... 45
12. Valores de ILD para os dados analisados com os diferentes programas..... 48

## RESUMO

A análise filogenética dos grupos basais de Diptera utilizando dados moleculares tem sido alvo de um número ainda limitado de trabalhos. Há um único trabalho propondo uma filogenia para os grupos basais de Diptera com esse tipo de informação até o presente momento. No presente trabalho, foram realizadas análises filogenéticas para grupos de Diptera, utilizando dois genes nucleares (18S e 28S) e quatro genes ribossomais (12S, 16S, COI e COII). As análises foram implementadas no programa POY 3.0.11 utilizando o método de Otimização Direta de Dados Moleculares (“Direct Optimization”) com diferentes matrizes de custos para *indels*, transições e tranversões, sendo utilizadas quatro mil réplicas por análise. Foram testados os efeitos do uso de diferentes programas de alinhamento (CLUSTALW, MALIGN e POY) sobre as topologias, bem como a influência da diminuição de táxons dos grupos externo e interno e da utilização de diferentes regiões de um mesmo gene. Para essa discussão dos protocolos de análise, foram utilizadas seqüências do gene 28S. As topologias geradas pelas análises dos genes ribossomais foram profundamente discordantes das hipóteses filogenéticas baseadas em diferentes tipos de dados morfológicos. Para todos esses genes mitocondriais a ordem Diptera resulta parafilética e a monofilia de seus principais grupos não é recuperada. Nas análises realizadas com os genes nucleares, Diptera tem sua monofilia recuperada, mas as topologias demonstram pontos de concordância e alguns pontos de discordância quanto a hipóteses de monofilia de grandes grupos, relações filogenéticas entre famílias dentro das subordens e relações filogenéticas entre as subordens. Nas análises realizadas para o gene 28S com as diferentes variáveis, pelo menos dez hipóteses diferentes de relação entre as subordens foram encontradas. Isso demonstra que ainda há problemas graves para obter topologias minimamente confiáveis utilizando dados moleculares. As causas dos conflitos entre diferentes topologias obtidas com dados moleculares e delas com dados de morfologia devem ser variadas e incluem: confiabilidade das seqüências originais; complexidade matemática subjacente ao problema de alinhamento e de análise de grandes matrizes; escassez de dados, resultando em subamostragem de grupos internos e externos e, por conseqüência, *long-branch attraction*; algoritmos impróprios para análise filogenética. É recomendável extrema cautela com reinterpretações evolutivas no momento, utilizando filogenias atualmente obtidas com dados moleculares.

## ABSTRACT

Phylogenetic analyses of the basal relationships of Diptera using molecular data have been performed a limited number of studies. Actually, there is a single paper proposed with a phylogenetic hypotheses for the basal groups of Diptera using this kind of information. In this work phylogenetic analyses of Diptera have been performed, using data from two nuclear genes (18S and 28S) e four mitochondrial genes (12S, 16S, COI and COII). The analyses were implemented in the POY program (version 3.1.11) with the Direct Optimization method, under different cost matrices for indels, transitions and transversions. The effects of the different alignment programs on the topologies (CLUSTALW, MALIGN and POY), as well as the influence of the reduction of number of taxa in the both ingroup and outgroup, and the use of different regions of the same gene (28S gene). The topologies generated with mitochondrial genes are highly incongruent with the phylogenetic hypotheses proposed based on different kinds of morphological data. For all that mitochondrial genes the order Diptera results paraphyletic, and the monophyly of this main clades is not recovered. In the analysis with the nuclear genes Diptera results monophyletic, but the topologies demonstrated only partially agree with the monophyly of this main clades, with the phylogenetic relationships between families within the suborders, and with phylogenetic relationships between the suborders. In the analyses performed for the 28S gene with the different parameters, at least ten different hypothesis of relationships among the suborders has found. These results demonstrate that it is still a challenge to obtain reliable topologies with molecular data. There are different reasons for the conflicts between the different topologies obtained with molecular data and between topologies obtained using molecular and morphological data: reliability of the original sequences; the mathematical underlying of alignment and analyses of large matrices; scarcity of data, resulting in poorly sampled ingroups and outgroups, with consequently long-branch attraction effect; and use of inappropriate algorithms for phylogenetic analyses. Extreme prudence is commendable while making evolutionary re-interpretations using topologies obtained using molecular data.

## 1. INTRODUÇÃO

### 1.1. Diversidade e Sistemática de Diptera

A ordem Diptera é um dos grupos de insetos com maior número de espécies, variação morfológica e diversidade ecológica. Há cerca de 125.000 espécies já descritas, mas estima-se que o número de espécies existentes seja de pelo menos o dobro (Yaetes & Wiegmann, 1999). Os dípteros vivos estão classificados, dependendo do autor, em pelo menos 130 famílias, 22-32 superfamílias, 8-10 infraordens e 5 subordens (McAlpine & Wood, 1989; Colless & McAlpine, 1991; Yaetes & Wiegmann, 1999, Amorim & Yaetes, 2006). O grupo teve origem no Permiano, com diversificação inicial no Triássico (Hennig, 1981; Kukalová-Peck, 1991; Amorim & Silva, 2002).

Várias sinapomorfias suportam a monofilia de Diptera (Wood & Borkent, 1989; Yaetes & Wiegmann, 1999), sendo também reconhecidamente monofilética a subordem Brachycera (Oosterbroek & Courtney, 1995; Friedrich & Tautz, 1997a). Existe discordância entre as hipóteses sobre qual seria o grupo irmão de Diptera. Hennig (1981) considera Mecoptera como táxon mais próximo, nomeando de Antliophora o clado formado por essas duas ordens. Posteriormente, Kristensen (1975) propôs que Mecoptera e Siphonaptera formariam um grupo monofilético, e esse clado seria o grupo irmão de Diptera. Wood & Borkent (1989) propuseram que somente a família Nannochoristidae, normalmente incluída em Mecoptera, seja grupo irmão de (Siphonaptera + Diptera). Mais recentemente, análises filogenéticas utilizando o gene ribossomal 18S (Whiting & Wheeler, 1994) e análises combinadas dos genes 18S e 28S (Whiting *et al.*, 1997) colocaram a ordem Strepsiptera dentro de Antliophora como grupo irmão de Diptera. Essa posição é questionada por alguns autores (Carmean & Crespi, 1995; Huelsenbeck, 1997) devido à possibilidade dessa união ser ocasionada por *long-branch attraction* (Felsenstein, 1978).

Estudos mais detalhados abordando sistemática de Diptera datam do início do século XIX (Meigen, 1800, 1803; Latreille, 1802 *apud* Friedrich & Tautz, 1997a). Os estudos de Willi Hennig, tanto na sistematização da metodologia de análise filogenética (1966), quanto nas relações entre os subgrupos de Diptera (1968, 1973, 1981) contribuíram muito para o entendimento da evolução do grupo. Um estudo amplo da filogenia de Diptera foi publicado no Volume 3 do *Manual of Nearctic Diptera* (1989), no qual são abordadas as relações entre os

grandes subgrupos de Diptera em três capítulos: grupos basais de Diptera “Nematocera” (Wood & Borkent, 1989), grupos basais de Brachycera (Woodley, 1989) e Cyclorrhapha (McAlpine, 1989). Após a publicação dessa obra, um número maior de outros estudos foi publicado. Com estudos baseados em morfologia, o trabalho de Wood & Borkent (1989) foi seguido por Griffiths (1990, 1994), Courtney (1990, 1991), Oosterbroek & Theowald (1991), Wood (1991), Sinclair (1992), Krzeminski (1992), Amorim & Carvalho (1992), Amorim (1993), Blaschke-Berthold (1994), Amorim (1995), Oosterbroek & Courtney (1995), Michelsen (1996), Amorim *et al.* (1996), Saether (2000) e Mazzarolo & Amorim (2000). Todos esses trabalhos convergem, em termos gerais, para cinco subgrupos maiores de Diptera: Tipulomorpha, Psychodomorpha, Bibionomorpha, Culicomorpha e Brachycera. Em alguns dos trabalhos, porém, Axymyiomorpha, Ptychopteromorpha e Blephariceromorpha são sugeridos como clados independentes. Entretanto, a composição desses subgrupos, bem como suas relações são consideravelmente diferentes entre as hipóteses desses autores. Segundo Yaetes & Wiegmann (1999), a plasticidade anatômica dos dípteros torna difícil, em alguns casos, o reconhecimento de homologia primária entre algumas estruturas, o que seria uma explicação para tanta controvérsia entre as hipóteses filogenéticas baseadas em morfologia.

## 1.2. Infraordens de Diptera

### 1.2.1. Tipulomorpha

Existem controvérsias entre os diversos autores sobre a composição desse grupo, bem como sobre a sua posição dentro de Diptera. Hennig (1973, 1981) inclui em Tipulomorpha as famílias Tipulidae (formada por Tipulidae, Limoniidae e Cylindrotomidae) e Trichoceridae, considerando esse grupo, denominado Polyneura, como grupo irmão de todos os outros dípteros. Essa classificação é discutida por Wood & Borkent (1989), que incluem em Tipulomorpha somente a família Tipulidae, mas também a posicionam na base da filogenia de Diptera, bem como Courtney (1991), Wood (1991) e Sinclair (1992). Krzeminski (1992) retoma a formação de Polyneura proposta por Hennig (Tipulidae + Trichoceridae), mas não o posiciona como grupo irmão dos outros dípteros. Tipulomorpha é posicionado por Krzeminski (1992) como grupo irmão de Ptychopteromorpha, e Tanyderidae é sugerido como grupo mais basal de Diptera. Griffiths (1990) e Oosterbroek & Theowald (1991) também não sustentam Tipuloidea (= Tipulomorpha *sensu* Wood & Borkent, 1989) como grupo irmão do

restante de Diptera. Griffiths (1990) sugere a inclusão de Tipuloidea próximo de Trichoceridae, dentro de Psychodomorpha. Oosterbroek & Courtney (1995) apresentam Tipulidae + Trichoceridae dentro dos chamados “higher Nematocera” formando um clado monofilético juntamente com Anisopodidae + Brachycera. Amorim *et al.* (1996) assumem que as diferenças de resultados entre os resultados de Oosterbroek & Courtney (1995) e os estudos anteriores (Wood & Borkent, 1989; Courtney, 1991; Sinclair, 1992) são devidas a problemas na análise desses autores, como a inserção de *missing data*, problemas de homologia primária, codificação, polarização e interpretações de homologias secundárias. Michelsen (1996) sugere um grupo monofilético na base de Diptera composto por Tipuloidea + Trichoceridae + Tanyderidae + Ptychopteridae.

### 1.2.2. Psychodomorpha

Hennig (1981) inclui dentro dessa infraordem as famílias Deuterophlebiidae, Blaphariceridae, Tanyderidae, Ptychopteridae e Psychodidae, e a posiciona na base de Oligoneura, que compreende todos os dípteros menos Tipulomorpha (ou Polyneura). A composição dessa infraordem proposta por Wood & Borkent (1989) difere consideravelmente da proposta por Hennig. Esses autores reúnem as famílias Trichoceridae, Perissomatidae, Anisopodidae, Scatopsidae e Canthyloscelidae (=Synneuridae) juntamente com Psychodidae; as famílias Ptychopteridae e Tanyderidae são reunidas em uma outra infraordem, Ptychopteromorpha. Krzeminski (1992) separa as famílias incluídas nesse grupo em diferentes clados, não assumindo essa infraordem. Além disso, o posicionamento filogenético desse grupo (*sensu* Wood & Borkent, 1989) dentro de Diptera difere entre autores. Courtney (1990, 1991) sugere duas hipóteses: como grupo irmão de Blephariceromorpha ou como grupo irmão de Ptychopteromorpha + Culicomorpha, sendo essa a mesma hipótese proposta por Wood & Borkent (1989). Outras hipóteses incluem Brachycera como grupo irmão de Psychodomorpha (Sinclair, 1992; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996). Michelsen (1996) mantém nessa infraordem apenas Trichoceridae e Psychodidae, incluindo as outras famílias no clado Neodiptera.

### 1.2.3. Bibionomorpha

Bibionomorpha é considerado por Hennig (1981) grupo irmão de Brachycera e as famílias incluídas por ele nesse grupo são: Canthyloscelidae, Scatopsidae, Cecidomyiidae,



Sciaridae, Mycetophilidae, Bibionidae (*Bibionomorpha sensu stricto*), Perisommatidae, Cramptonomyiidae, Pachyneuridae, Axymyiidae e Anisopodidae (*Bibionomorpha sensu lato*). Wood & Borkent (1989) retiram de *Bibionomorpha* as famílias Canthyloscelidae, Perisommatidae, Anisopodidae e Scatopsidae, transferidas para *Psychodomorpha*, e Axymyiidae, grupo monofilético de *Axymyiomorpha*. As relações filogenéticas dessa infraordem dentro de *Diptera* não são esclarecidas por Wood & Borkent (1989), já que o cladograma sugerido por esses autores é apenas parcialmente resolvido, apresentando uma politomia incluindo os clados *Bibionomorpha*, *Axymyiomorpha*, *Blephariceromorpha* e *Psychodomorpha* + (*Ptychopteromorpha* + *Culicomorpha*). Courtney (1990, 1991) e Sinclair (1992) consideram *Bibionomorpha* como grupo irmão do restante de *Diptera*, menos *Tipulomorpha*. Hennig (1981) e Amorim (1993) posicionam *Bibionomorpha* como grupo irmão de *Brachycera*. O mesmo é considerado por Michelsen (1996), que denomina o clado formado por *Bibionomorpha s.l.* + *Brachycera* como *Neodiptera*.

A filogenia de *Bibionomorpha* foi estudada em detalhe por Amorim & Rindal (no prelo), com *Anisopodidae* como grupo basal, seguido de *Pachyneuridae*, *Bibionidae* e os *Mycetophiliformia*, composto por *Cecidomyiidae*, *Sciaridae*, *Rangomaramidae*, os *Mycetophiloidea* e os *Keroplastidae*.

#### **1.2.4. Culicomorpha**

*Culicomorpha* é um clado amplamente aceito como monofilético e é dividido em duas superfamílias: *Culicoidea*, que inclui as famílias *Culicidae*, *Dixidae*, *Corethrellidae* e *Chaoboridae*; e *Chironomoidea*, incluindo as famílias *Thaumaleidae*, *Simuliidae*, *Ceratopogonidae* e *Chironomidae* (Saether, 2000). Não existem muitas hipóteses discordantes sobre o posicionamento filogenético de *Culicomorpha* dentro de *Diptera*, sendo que a maioria dos estudos converge para a hipótese de monofilia do clado *Culicomorpha* + *Ptychopteromorpha* (Wood & Borkent, 1989; Courtney, 1990, 1991; Sinclair, 1992; Oosterbroek & Courtney, 1995). No entanto, outras hipóteses são sugeridas, como *Culicomorpha* + *Bibionomorpha s.l.* (Hennig, 1981); *Culicomorpha* + ((*Anisopodomorpha* + *Bibionomorpha*) + *Brachycera*) (Krzeminski, 1992); *Culicomorpha* + (*Psychodomorpha* + *Brachycera*) (Amorim *et al.*, 1996); ou a família *Thaumaleidae* + *Culicomorpha* (Saether, 2000).

#### **1.2.5. Axymyiomorpha**

A família Axymiidae é posicionada por Hennig (1981) e Oosterbroek & Courtney (1995) dentro da infraordem Bibionomorpha. No entanto, outros autores a consideram como uma infraordem independente (Wood & Borkent, 1989; Griffiths, 1990; Courtney, 1990, 1991 e Sinclair, 1992).

#### **1.2.6. Ptychopteromorpha**

A infraordem Ptychopteromorpha inclui as famílias Ptychopteridae e Tanyderidae. Essa infraordem é freqüentemente considerada como grupo irmão de Culicomorpha (Wood & Borkent, 1989; Courtney, 1990, 1991; Sinclair, 1992; Oosterbroek & Courtney, 1995). No entanto, essas famílias também aparecem dentro de Psychodomorpha (Hennig, 1981; Amorim *et al.*, 1996).

#### **1.2.7. Blephariceromorpha**

Três famílias são incluídas nessa infraordem, Blephariceridae, Deuterophlebiidae e Nymphomyiidae. Courtney (1990, 1991) sugere que Blephariceromorpha seja grupo irmão de Psychodomorpha ou de Psychodomorpha + (Ptychopteromorpha + Culicomorpha). Outra hipótese é sugerida, onde essa infraordem é sugerida como grupo irmão de Bibionomorpha + (“higher Nematocera” + Brachycera) (Oosterbroek & Courtney, 1995).

#### **1.2.8. Brachycera**

Brachycera é certamente um grupo monofilético, suportado por um grande número de sinapomorfias (Wood & Borkent, 1989; Sinclair, 1992; Oosterbroek & Courtney, 1995; Friedrich & Tautz, 1997a; Yaetes & Wiegmann, 1999). As relações entre as infraordens de Brachycera permanecem incertas (Griffiths, 1994; Hennig, 1981; Woodley, 1989), havendo consenso quanto à monofilia de Cyclorrhapha e (Empidoidea + Cyclorrhapha).

#### **Xylophagomorpha**

Muitos autores consideram essa infraordem composta por apenas uma família, Xylophagidae (Griffiths, 1994; Woodley, 1989). Griffiths (1994) sugere a união de Xylophagomorpha e Tabanomorpha com base em sinapomorfias da genitália do macho, mas outros autores sugerem que a hipótese de monofilia desse clado não é robusta (Yaetes & Wiegmann, 1999).

### Tabanomorpha

Cinco famílias são incluídas nesse grupo: Tabanidae, Pelecorhynchidae, Rhagionidae, Athericidae e Vermileonidae, embora a posição dessa última família seja incerta e as vezes colocada em uma infraordem própria (Sinclair, 1992; Woodley, 1989; Yaetes & Wiegmann, 1999). A monofilia de Rhagionidae é bastante debatida.

### Stratiomyomorpha

Essa infraordem contém duas famílias, Xylomyidae e Stratiomyidae, sendo que Pantophtalmidae também é incluída nesse grupo por alguns autores (Griffiths, 1990; Sinclair, 1992, Mazzarolo & Amorim, 2000) como grupo irmão de Xylomyidae + Stratiomyidae.

### Muscomorpha

Woodley (1989) aumenta a abrangência dessa infraordem sobre a apresentada por McAlpine *et al.* (1981) no Volume 1 do *Manual of Nearctic Diptera*, incluindo a infraordem Asilomorpha, exceto Vermileonidae. Assim, Muscomorpha incluiria todas as famílias de Diptera exceto aquelas incluídas nas infraordens discutidas acima.

### Nemestrinoidea

Considerada como grupo irmão de Heterodactyla (que inclui todos os Muscomorpha menos Nemestrinoidea), esse grupo não tem sua monofilia muito bem sustentada, mas as duas famílias incluídas nesse grupo são consideradas monofiléticas: Nemestrinidae e Acroceridae (Sinclair, 1992; Woodley, 1989; Yaetes & Wiegmann, 1999).

### Asiloidea

As famílias Asilidae, Apioceridae, Mydidae, Scenopinidae, Therevidae e Bombyliidae são incluídas em Asiloidea. Woodley (1989) posiciona Bombyliidae como grupo irmão do restante de Asiloidea, enquanto Yeates (1994) posiciona Bombyliidae + Hilarimorphidae na base desse grupo. Algumas famílias desse grupo parecem não ser monofiléticas. A monofilia de Bombyliidae não é bem suportada, sendo que pode ser dividida em quatro famílias (Yeates, 1994). Therevidae e Apioceridae também são consideradas parafiléticas (Yeates, 1994, Yeates & Wiegmann, 1999).

## Empidoidea

O clado formado pelas famílias Empididae e Dolichopodidae é considerado monofilético (Woodley, 1989; Sinclair, 1992; Wiegmann *et al.*, 1993; Griffiths, 1996). Em recente trabalho, Sinclair & Cumming (2006) confirmam a monofilia de Empidoidea, bem como de Empidoidea + Cyclorrhapha (Eremoneura). Esses autores incluem em Empidoidea as famílias Hybotidae, Atelestidae e Brachystomatidae. No entanto, as relações entre essas famílias são parcialmente resolvidas.

## Cyclorrhapha

A infraordem Cyclorrhapha (ou Muscomorpha *sensu* McAlpine, 1989) inclui um grande número de famílias e com grande plasticidade morfológica (Yeates & Wiegmann, 1999). A monofilia desse clado é bem suportada (McAlpine, 1989). Essa infraordem é dividida em dois subgrupos: Aschiza e Schizophora. Aschiza, que compreende as famílias Platypezidae, Lonchopteridae, Ironomyiidae, Sciadoceridae, Phoridae, Syrphidae, Pipunculidae e Opetidae, é considerado por McAlpine (1989) como um grupo monofilético. No entanto, alguns autores consideram esse subgrupo parafilético com relação à Schizophora (Griffiths, 1990). Schizophora compreende a maior diversidade de famílias de Diptera (McAlpine, 1989; Yeates & Wiegmann, 1999) e é dividido em dois grandes grupos: Acalyptratae e Calyptratae.

### 1.3. Estudos moleculares em Diptera

Em Diptera, a maioria dos estudos filogenéticos com dados moleculares aborda as relações entre pequenos grupos. O estudo de relações entre espécies do gênero *Drosophila* é abundante na literatura, sendo abordadas, principalmente, relações em grupos de espécies (Suzuki *et al.*, 1990; Pelandakis *et al.*, 1991; Barrio *et al.*, 1992; Barrio *et al.*, 1994; Caccone *et al.*, 1996; Nikolaidis & Scouras, 1996; Nurminsky *et al.*, 1996; Spicer & Pitnick, 1996; Piano *et al.*, 1997; O'Grady *et al.*, 1998; O'Grady, 1999; Durando *et al.*, 2000; O'Grady & Kidwell, 2002).

Diferentes gêneros também têm tido suas relações filogenéticas investigadas a partir de dados moleculares, como *Chironomus* (Chironomidae) (Guryev *et al.*, 2001), *Phytomyza* (Agromyzidae) (Scheffer & Wiegmann, 2000), *Rhagoletis* (Tephritidae) (Smith & Bush, 1997), *Bactrocera* (Tephritidae) (Muraji & Nakahara, 2001; Smith *et al.*, 2003), *Cheilosia* (Syrphidae) (Stahls & Nyblom, 2000), *Boettcherisca* (Sarcophagidae) (Kurahashi & Kano, 1984), *Tricops* (Muscidae) (Savage *et al.*, 2004), *Ophyra* (Muscidae) (Schuehli *et al.*, 2004), da mesma forma que

gêneros de importância médica, como *Anopheles* (Culicidae) (Besansky *et al.*, 1994; Porter & Collins, 1996; Foley *et al.*, 1998; Chen *et al.*, 2002; Krzywinski & Besansky, 2003; Sallum *et al.*, 2002), *Phlebotomus* (Psychodidae) (Depaquit *et al.*, 2000; Di Muccio *et al.*, 2000) e *Lutzomyia* (Psychodidae) (Ready *et al.*, 1998; Ishikawa *et al.*, 1999).

Pequenos grupos de espécies de Diptera foram utilizados como estudo de caso para testes de utilidade de certos genes para análises filogenéticas, como o gene mitocondrial citocromo oxidase II (COII) (Pruess *et al.*, 2000; Baker *et al.*, 2001), 12S, 16S, fator de alongamento 1 $\alpha$  (EF-1 $\alpha$ ), *wingless* e *white* (Baker *et al.*, 2001). Alguns autores abordam as discrepâncias entre hipóteses geradas por dados morfológicos e moleculares (Thomas & Hunt, 1993; Kwiatowski & Ayala, 1999), enquanto outros preferem, ao contrário, combinar esses dois tipos de dados para a elaboração de suas hipóteses (Meier & Wiegmann, 2002).

Estudos das relações filogenéticas entre grupos mais abrangentes - como famílias - de Diptera são mais raros, sendo poucos os que incluem um número representativo dos gêneros de uma família. Entre esses, pode-se citar DeSalle (1992, para Drosophilidae), Han & McPherson (1997, para Tephritidae), Bernasconi *et al.* (2000b, para Scathophagidae), Meier & Wiegmann (2002, para Coelopidae) e Cook *et al.* (2004, para Phoridae). Ainda mais escassas são as hipóteses envolvendo grandes grupos de dípteros. Para essas análises mais abrangentes, há os estudos de Culicomorpha (Miller *et al.*, 1997), Tabanomorpha (Wiegmann *et al.*, 2000), Calyptratae (Bernasconi *et al.*, 2000a; Nirmala *et al.*, 2001), Empidoidea (Collins & Wiegmann, 2002a) e Cyclorrhapha (Collins & Wiegmann, 2002b).

Até o momento, a única hipótese filogenética para os grupos basais da ordem Diptera baseada em dados moleculares foi elaborada por Friedrich & Tautz (1997a), utilizado fragmentos do gene nuclear 28S. No entanto, esse estudo baseou-se ainda em um conjunto muito pequeno de táxons terminais (somente 16 espécies representando todos os grandes subgrupos), resultando em uma filogenia pouco resolvida para as relações entre os principais subgrupos de Diptera. Esses autores obtiveram como clados monofiléticos em seu trabalho Culicomorpha, Tipulomorpha (menos Trichoceridae), Brachycera e Bibionomorpha (incluindo Scatopsidae).

#### **1.4. Análise de dados moleculares**

O crescimento recente no número de estudos de filogenia baseados em dados moleculares possivelmente pode ser atribuído a duas causas básicas: o advento de uma tecnologia para obtenção rápida e relativamente barata de dados moleculares e o desenvolvimento de métodos rigorosos para análises filogenéticas (Sanderson & Shaffer, 2002). Dados moleculares têm sido extensivamente usados não somente para procurar esclarecer relações filogenéticas difíceis de abordar por outros métodos, mas também para questionar hipóteses filogenéticas propostas com base em outras fontes de caracteres (Russo, 2001).

Atualmente, a quantidade de seqüências de genes de todos os organismos disponíveis em bancos de dados é gigantesca. Essas seqüências são de genes nucleares, genomas organelares (mitocôndrias e cloroplastos) ou de fragmentos. Apenas para a ordem Diptera, em abril de 2006 estavam depositadas no GenBank mais de 2.000.000 de referências. O esforço em gerar seqüências de genes, por outro lado, tem sido desproporcionalmente maior que o esforço em analisar em profundidade e extensão as seqüências já disponíveis. Além disso, o aumento da facilidade e o decréscimo de custos para a geração de grande quantidade de dados moleculares resultam em um aumento na demanda computacional para análises filogenéticas, tanto para alinhamentos quanto para reconstrução de topologias (Terry & Whiting, 2005).

Do ponto de vista metodológico, a reconstrução filogenética com a utilização de dados moleculares apresenta uma diversidade de problemas específicos, listados e discutidos por Sanderson & Shaffer (2002). Isso inclui a geração de seqüências confiáveis e todo o processo de análise, desde homologia entre as seqüências e alinhamento, até o tempo computacional necessário para o processamento da análise.

Seqüências de DNA que não apresentam o mesmo tamanho, ou seja, o mesmo número de bases, necessitam sofrer um processo de alinhamento antes mesmo do início da análise filogenética, isto é, da reconstrução da topologia. O alinhamento é o processo no qual os nucleotídeos são individualizados como diferentes caracteres e há a decisão de quais nucleotídeos de cada uma das seqüências estudadas podem ser tomados como primariamente homólogos. Isso implica em transformar diversas seqüências de DNA em uma matriz de dados, onde cada nucleotídeo (A, C, G ou T) é um estado de caráter. Para que as seqüências apresentem, ao final do processo, o mesmo número de bases ou caracteres, torna-se necessária a inclusão de *indels* - hipóteses de deleção ou inserção de uma base para uma determinada posição. Diferentes alinhamentos são obtidos à medida que *indels* são distribuídos nas

seqüências, alterando o tipo e frequência das bases desemparelhadas. Os diferentes alinhamentos criam matrizes de caracteres distintas, alterando o resultado final.

A homologia de posição em seqüências ortólogas pode se apresentar como estados idênticos de caráter (nucleotídeos) em todas as seqüências, por substituição em uma base (representando mutações pontuais) ou uma inserção ou deleção (*indels*). O aspecto mais problemático do alinhamento de seqüências é o posicionamento de *indels* e esse problema aumenta com o aumento da divergência entre os táxons (Morrison & Ellis, 1997).

Atualmente, vários programas são utilizados para proceder ao alinhamento de seqüências de DNA. Esses programas utilizam três premissas distintas para a realização dos alinhamentos: alinhamento múltiplo de seqüências (Higgins *et al.*, 1996; Thompson *et al.*, 1994, 1997), otimização direta (Wheeler, 1996) e *search-based optimization* (Wheeler, 1999, 2003). O alinhamento múltiplo de seqüências precede a construção do cladograma, enquanto que nos outros dois métodos o processo de alinhamento é feito em etapas consecutivas junto com a construção de cladogramas (Wheeler, 2001), em um processo de iluminação recíproca.

O alinhamento múltiplo de seqüências é realizado para tornar seqüências de tamanhos diferentes em um número igual de caracteres através de inserção de *indels*. Os três programas para alinhamentos múltiplos mais usados hoje em dia são CLUSTAL (Higgins & Sharp, 1988; Higgins *et al.*, 1992, 1996; Higgins, 1994; Jeanmougin *et al.*, 1998; Thompson *et al.*, 1994, 1997), TREEALIGN (Hein, 1989, 1990) e MALIGN (Wheeler & Gladstein, 1994). Cada um desses métodos conta com uma árvore guia (*guide tree*) para direcionar a construção do alinhamento. Phillips *et al.* (2000) constataram que na maioria das análises estáticas a topologia resultante da análise filogenética difere da árvore guia usada para gerar o alinhamento. Essa observação sugere que a análise foi realizada com base em um alinhamento subótimo (menos parcimonioso).

Com o CLUSTAL e o TREEALIGN, uma árvore de distância é calculada para os custos do alinhamento de pares. Para cada nó das árvores guias, uma seqüência de consenso (CLUSTAL) ou quase-otimizada (TREEALIGN) é criada para o par alinhado e é submetida a outro pareamento de pares, promovendo a construção da árvore. Quando a raiz da árvore é alcançada, os vários *indels* inseridos criam as seqüências de mesmo tamanho, ou seja, o múltiplo alinhamento (Phillips *et al.*, 2000; Wheeler, 2001).

O programa MALIGN também cria árvores guias, mas difere dos outros dois programas porque examina diversas árvores, criadas paralelamente por procedimentos de

construção de árvores e *branch swapping*. Durante o procedimento de busca, um alinhamento múltiplo completo é gerado para cada árvore guia e uma análise filogenética heurística é feita para os alinhamentos múltiplos. O alinhamento que produzir o resultado filogenético mais parcimonioso é escolhido como o “melhor” alinhamento (Wheeler & Gladstein, 1994; Phillips *et al.*, 2000; Wheeler, 2001).

Esses três programas de alinhamento geram o mesmo tipo de resultados: uma matriz de caracteres, que é submetida à análise filogenética. Quando a análise da matriz é feita, as supostas homologias não são alteradas ou reexaminadas. As homologias são criadas *a priori* e não são revisadas depois dessa etapa. Dessa forma, esse tipo de análise é denominada de **estática** (Wheeler, 2001).

O método de otimização direta cria um único grupo de supostas homologias para cada topologia. Segundo Wheeler (2001), com esse método, são minimizados os eventos de mudanças de nucleotídeos e inserção de *indels*, sendo que correspondências entre bases minimizam o comprimento da árvore. Do mesmo modo como no alinhamento múltiplo, as homologias são criadas base a base, entretanto, essas homologias são dinamicamente determinadas, adaptadas precisamente para cada topologia. O procedimento é baseado em uma determinação heurística de seqüências ancestrais hipotéticas. O algoritmo começa com um nó interno com dois terminais. A seqüência do ancestral hipotético mais parcimoniosa é criada para esse nó via um procedimento dinâmico, onde todas as seqüências ancestrais hipotéticas são implicitamente examinadas. Isso é executado de uma forma muito semelhante ao alinhamento de pares. Esse método não requer um prévio alinhamento múltiplo de seqüências (Wheeler, 1996; Wheeler, 2001; Giribet *et al.*, 2002; Terry & Whiting, 2005).

A vantagem desse método é que ele permite literalmente a avaliação de milhões de topologias com árvores guias alternativas. Ele é logicamente consistente, no sentido de que o mesmo critério é utilizado para construir o alinhamento e a topologia final, permitindo uma otimização dos dados. A desvantagem encontra-se no complexo cálculo associado com a avaliação de todas as topologias e de todas as formas de otimizar o grupo de seqüências para cada topologia. A busca da topologia ótima por POY é um problema *NP-complete*, ou seja, um problema matemático para o qual nenhum algoritmo efetivo é conhecido. O aumento do número de táxons terminais aumenta o número de possibilidades a serem examinadas, levando a um aumento astronômico do tempo computacional, inexecutável para os computadores atuais. O fato de um problema ser um *NP-complete* é considerado justificativa para a utilização



de procedimentos heurísticos, isto é, procedimentos que não garantem produzir uma solução “ótima” para todos os casos do problema (Graham & Foulds, 1982). Para resolver esse problema, o programa utiliza várias opções de refinamento, que permitem uma busca bastante rigorosa. Dada uma solução ótima, o POY pode gerar um alinhamento implícito, semelhante à um alinhamento múltiplo, mas com os estados de caráter unindo estados ancestrais e descendentes (Wheeler, 2003; Terry & Whiting, 2005).

Mesmo com o método de otimização direta, seqüências de comprimentos muito diferentes podem ser muito difíceis de analisar. Pode acontecer de alguns terminais do grupo em análise possuírem seqüências com centenas de nucleotídeos, enquanto outros possuam poucos ou nenhum. A inserção de *indels* pelo programa, nesses casos, pode ser virtualmente casual, prejudicando a análise. Isso se deve ao fato de que cada nucleotídeo é tratado como uma potencial homologia. O método de *Search-based optimization* (Wheeler, 1999, 2003) trata cadeias contínuas de nucleotídeos como estados de caráter de um caráter complexo. Os custos de transformação entre os estados correspondem à soma dos vários eventos requeridos para transformar uma seqüência em outra. Uma vez que o grupo de estados potenciais é criado, a análise é procedida por otimização dinâmica.

Além do alinhamento, que é a base para a geração dos cladogramas, outro fator que influencia a topologia é a natureza - molecular ou morfológica - de dados utilizados nas análises. Quando é utilizado mais de um conjunto de caracteres em uma análise, pode haver conflitos entre as diversas fontes, sendo que cada análise pode resultar em topologias discordantes. Kluge (1989) sugere aumentar a independência de caracteres, obtendo-se diferentes classes ou tipos, como dados morfológicos e moleculares. Esses diferentes conjuntos de dados podem ser combinados em uma única análise (Evidência total ou Análise simultânea), e testados separadamente para verificar a presença de conflitos. Uma forma de medir se os dados utilizados em uma análise são conflitantes ou não é fazendo uma análise de congruência. A primeira medida de congruência utilizada em análises de sensibilidade foi o ILD (*incongruence length difference*). O ILD foi originalmente desenvolvido como uma medida de homoplasias entre conjuntos de dados, quando esses eram combinados em uma análise (Mickey & Farris, 1981). Posteriormente, foi implementado como um teste estatístico para determinar se conjuntos de dados poderiam ser utilizados em análises combinadas (Swofford, 2002). Atualmente, o ILD é utilizado como um critério para selecionar parâmetros de alinhamentos procurando minimizar incongruências entre partições de dados (Wheeler &

Hayashi, 1998; Terry & Whiting, 2005), ou seja, procurando reconhecer os sinais filogenéticos e suprimir os ruídos (Meier & Wiegmann, 2002). O ILD é calculado baseando-se na diferença de comprimento de árvores entre a árvore resultante da análise combinada dos dados e a soma do comprimento das árvores de cada partição:

$$ILD = \frac{L(x+y) - (L_x + L_y)}{L(x+y)},$$

onde  $L(x+y)$  é o comprimento total da árvore com os dados combinados, e  $L_x$  e  $L_y$  são os comprimentos da árvore mais parcimoniosa de cada partição, separadamente.

Essas diferenças de procedimento seriam pouco relevantes se chegassem, por vias distintas, aos mesmos resultados. Contudo, não é isso que ocorre. Com a existência de um número gigantesco de alinhamentos possíveis e de árvores possíveis em análises filogenéticas (Felsenstein, 1982), as diferenças de procedimento resultam em diferenças enormes entre topologias resultantes, como será visto adiante. Topologias distintas implicam em interpretações diferentes (às vezes incompatíveis) de evolução de características e de evolução biogeográfica. Desse modo, a escolha de protocolos adequados de análise em absoluto não é uma questão trivial em análises filogenéticas com dados moleculares.

## **2. OBJETIVOS**

O objetivo dessa tese é elaborar hipóteses de relações filogenéticas entre as famílias basais de Diptera, utilizando dados provenientes de seqüências de DNA - genes nucleares ribossomais 18S e 28S, e dos genes mitocondriais 12S, 16S, citocromo oxidase subunidade I e citocromo oxidase subunidade II - com um protocolo de análise que permita verificar a robustez dos dados, considerando a questão da amostragem de táxons terminais e dos métodos de alinhamento obtidos entre as seqüências.

## **3. MATERIAL E MÉTODOS**

### **3.1. Fontes das seqüências**

As seqüências dos genes nucleares 18S e 28S e dos genes mitocondriais citocromo oxidase I (COI), citocromo oxidase II (COII), 12S e 16S de diversas espécies e famílias da ordem Diptera foram obtidas no GenBank (<http://www.ncbi.nlm.nih.gov>). Os terminais utilizados em cada análise, bem como seus respectivos números de acesso no GenBank estão listados nos Anexos I a X.

### 3.2. Método geral de alinhamento

As seqüências dos táxons obtidas para cada gene foram pré-alinhadas com o programa de múltiplo alinhamento CLUSTAL (CLUSTAL W, versão 1.7; Thompson *et al.*, 1994) para a escolha dos táxons a serem utilizados. Foram selecionadas as seqüências que apresentaram alinhamento com regiões mais conservadas, ou seja, sem ambigüidades, buscando minimizar os erros de homologia primária entre as seqüências obtidas. Depois de selecionadas, essas foram re-alinhadas e os fragmentos mais conservados foram selecionados para a análise. Os comprimentos das seqüências utilizadas para os diferentes genes estão apresentados na Tabela 1.

**Tabela 1.** Número de pares de bases das seqüências utilizadas nas análises, antes dos alinhamentos.

Gene	Número de pares de bases
12S primeiro grupo	239 a 259
12S segundo grupo	287 a 301
16S primeiro grupo	356 a 372
16S segundo grupo	290 a 318
18S	357 a 482
28S	262 a 308
COI primeiro grupo	318
COI segundo grupo	256
COII	234

As seqüências dos genes 12S, 16S, 18S e 28S foram analisadas separadamente pelo método de otimização direta de dados moleculares com o programa POY 3.11 (Wheeler *et al.*, 2003). As análises foram realizadas no cluster do Centro de Insetos Sociais da Universidade Estadual de São Paulo – Rio Claro (CEIS UNESP/RC). O cluster tem 15 nós com processadores Athlon XP 2800+ (2GHz e 512 RAM) conectados em paralelo sob gerenciamento de PVM (Parallel Virtual Machine). Cada análise produziu topologias e alinhamentos implícitos segundo os comandos: ‘parallel –multirandom -

norandomizeoutgroup -nodiscrepancies -buildsperreplicate 10 -maxtrees 3 -replicates 1000 -fitchtrees -multidrft -driftspr -drifttbr -multiratchet -ratchettbr 5 -impliedalignment -molecularmatrix [nome do arquivo] -exact'. Foram utilizadas 10 matrizes de custos para *indels* : transversões : transições (Tabela 2), sendo os valores para cada tipo de mutação estabelecidos respeitando os limites apresentados por Wheeler (1995).

Para os genes 18S e 28S, também foram realizadas análises combinadas dos dois genes. Em virtude da ausência de informação referente às regiões utilizadas para alguns táxons terminais, foram realizados dois tipos de análises: (1) utilizando somente os táxons dos quais existem seqüências de DNA presentes para ambos os genes e (2) utilizando todos os táxons terminais, incluindo os que não tinham informações sobre as regiões em questão (com conseqüente inclusão de grande número de *missing data*).

**Tabela 2.** Custos para *indels*, transversões e transições utilizados nas análises dos genes 12S, 16S, 18S e 28S.

Parâmetro	<i>Indels</i>	Transversões	Transições
111	1	1	1
112	1	1	2
211	2	1	1
212	2	1	2
221	2	2	1
411	4	1	1
412	4	1	2
421	4	2	1
812	8	1	2
821	8	2	1

Para o gene COI foram testados 15 parâmetros de transversão:transição: 1:1; 2:0; 2:4; 2:6; 2:8; 4:0; 4:2; 4:6; 6:0; 6:2; 6:4; 6:8; 8:0; 8:2; 8:6. Para o gene COII foram testados 12 desses 15 parâmetros, excluindo os parâmetros 8:0; 8:2; 8:6. Não foram atribuídos custos para *indels* por essas seqüências não apresentarem inserção ou deleção de bases. Para cada parâmetro de custos, foram calculadas as árvores filogenéticas para cada um dos genes (COI e COII). Além das análises realizadas utilizando-se as seqüências completas, essas também foram particionadas da seguinte forma: primeiramente, foram desativadas as bases das segundas e terceiras posições de cada códon, calculando-se a árvore utilizando na matriz somente as bases das primeiras posições. Em seguida, foram ativadas as bases das segundas posições, desativando-se as bases das primeiras e terceiras posições, e finalmente, foram ativadas

somente as bases das terceiras posições dos códons. Dessa forma, foram obtidas árvores para três partições dos dados de cada um dos dois genes. Foi calculado o ILD para cada parâmetro de custos.

Essas análises foram feitas no programa 'TNT' (Hovenkamp, 2005), já que as seqüências desses genes mitocondriais não necessitam de alinhamento prévio. Os comandos utilizados nessa análise foram desenvolvidos pelo Prof. Dr. Fernando P. L. Marques e encontram-se no Anexo XI.

O enraizamento foi feito pelo método de comparação com grupos externos (Nixon & Carpenter, 1993). Foram utilizadas seqüências de espécies pertencentes a clados de diferentes níveis externos de Diptera - ordens Mecoptera, Siphonaptera, Lepidoptera, Trichoptera e Neuroptera e Coleoptera - visando diminuir a influência de homoplasias entre grupos externos e internos e problemas de otimização na base do grupo interno na filogenia. No entanto, existe na literatura a sugestão de que o grupo irmão de Diptera seja a ordem Strepsiptera (Whiting & Wheeler, 1994; Whiting *et al.*, 1997). Para testar essa hipótese, incluímos na matriz de dados para o gene 28S uma seqüência de *Stylops mellitae* (Strepsiptera, número de acesso Z98170), já que para os outros genes não existem seqüências para espécies dessa ordem.

### 3. 3. O gene nuclear 28S

As seqüências do gene 28S utilizadas por Friedrich & Tautz (1997a) foram alinhadas com outras seqüências de espécies de Diptera obtidas no GenBank por meio do programa BioEdit (Hall, 1999). Foram selecionadas e utilizadas nas análises as seqüências que não apresentaram ambigüidades, ou seja, as que apresentaram homologia com aquelas referidas acima, minimizando o risco de introdução de ruído desnecessário na análise e dificuldades na comparação com os resultados originais de Friedrich & Tautz (1997a).

As seqüências da região D7 do gene 28S foram alinhadas por três diferentes programas de alinhamento: CLUSTAL (Thompson *et al.*, 1994), MALIGN (Wheeler & Gladstein, 1994) e POY (Wheeler *et al.*, 2003). Os custos para *indels*, transversões e transições utilizados para os alinhamentos foram 1:1:1 para os três programas.

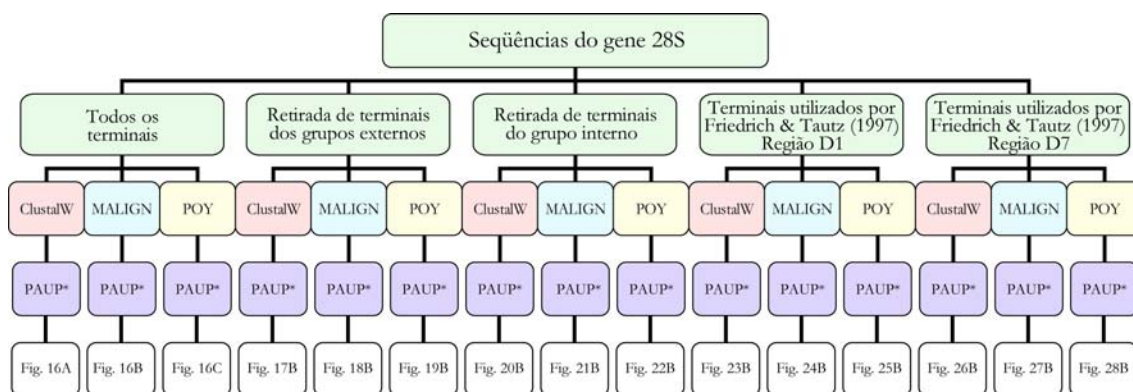
Os alinhamentos resultantes de cada programa foram analisados no programa PAUP\* (Swofford, 2000), sendo realizada uma busca heurística com *Branch and Bound*. No caso dos resultados do POY, foram utilizados os alinhamentos implícitos gerados pelos comandos: -

norandomizeoutgroup -molecularmatrix nomedoarquivo.txt -maxtrees 3 -replicates 100 -buildsperreplicate 10 -impliedalignment -exact.

As seqüências já alinhadas do gene 28S foram analisadas de diferentes formas. Inicialmente, foram feitas análises com todos os táxons selecionados (Anexo X). Em seguida, foram retirados terminais dos grupos externos (Trichoptera, Mecoptera e Siphonaptera), deixando-se apenas aqueles utilizados por Friedrich & Tautz (1997a), para a análise do efeito dessa diminuição no enraizamento dos cladogramas. Depois, mantendo o número inicial de terminais nos grupos externos, foram retirados terminais do grupo interno, mais especificamente, táxons da subordem Culicomorpha, para a verificação do efeito sobre a estabilidade dentro do clado. Finalmente, foram realizadas as análises deixando-se apenas os mesmos grupos utilizados por Friedrich & Tautz (1997a) (Figura 1). Além disso, as regiões D1 e D7 do gene 28S foram analisadas separadamente e em conjunto, para verificar o efeito do aumento do número de caracteres na matriz sobre a topologia final (Figuras 1 e 2).

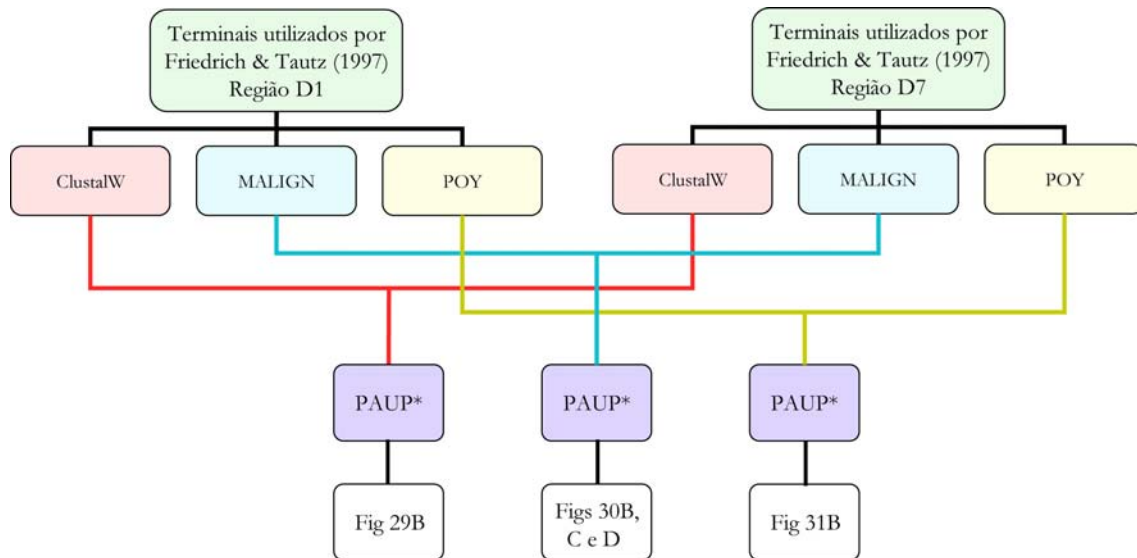
Os alinhamentos gerados pelos programas não sofreram qualquer tipo de ajuste manual *a posteriori*, indispensável para a discussão sobre a repetibilidade dos resultados. Esses ajustes demandam decisões subjetivas e conferem uma modificação nos parâmetros de custos utilizados no alinhamento final.

*Indels* foram tratados como quinto estado de caráter, por ser logicamente inconsistente tratá-los como *missing data*, já que *indels* correspondem a nucleotídeos anteriormente existentes ou inseridos e são interpretados como homologies nas análises filogenéticas (Giribet & Wheeler, 1999; Phillips *et al.* 2000).



**Figura 1.** Fluxograma das análises realizadas com os três diferentes programas de alinhamento.

Os alinhamentos implícitos gerados por POY para a região D7 do gene 28S, incluindo todos os táxons terminais, também foi analisado sob diferentes algoritmos: Neighbor-joining, Maximum Likelihood e Máxima Parcimônia. As três análises foram feitas no programa PAUP\* com os comandos de *default*. Para que essas análises pudessem ser realizadas, *indels* foram tratados como *missing data*, já que o tratamento como quinto estado de caráter não é aceito para Neighbor-joining nem para Maximum Likelihood.



**Figura 2.** Fluxograma representando a análise combinada das seqüências das regiões D1 e D7 do gene 28S, utilizando diferentes programas de alinhamento.

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Gene mitocondrial ribossomal 12S

A grande quantidade de seqüências disponíveis para o gene 12S permitiu que fossem definidos dois grandes grupos para as análises, sendo: (1) um grupo contendo 77 terminais no grupo interno, representando 23 famílias de Díptera, incluindo cinco famílias de dípteros basais (“Nematocera”) e 18 famílias de Brachycera (Anexo I); e (2) um grupo contendo 32 terminais no grupo interno, representando 17 famílias de Díptera, incluindo seis famílias de dípteros basais e 11 famílias de Brachycera (Anexo II).

#### 4.1.1. Primeiro grupo de terminais para o gene mitocondrial ribossomal 12S

Os dados das matrizes obtidas com o alinhamento implícito de POY sob diferentes parâmetros de custos para *indels*, transversões e transições, os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados estão sumarizados na Tabela 3.

**Tabela 3.** Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o primeiro grupo de terminais para o gene 12S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.

Parâmetro	Nº caracteres na matriz	Nº caracteres constantes	Nº caracteres informativos	Nº caracteres não informativos	Comprimento da árvore	Nº de MPT
111	484	55	307	122	1963	6
112	574	49	343	182	2338	7
211	417	54	260	103	2242	12
212	424	55	269	100	2701	4
221	475	55	294	126	3362	6
411	407	52	258	97	2691	12
412	419	51	259	109	3167	6
421	406	54	256	96	3903	3
812	386	42	263	81	3939	7
821	401	47	259	95	4797	3

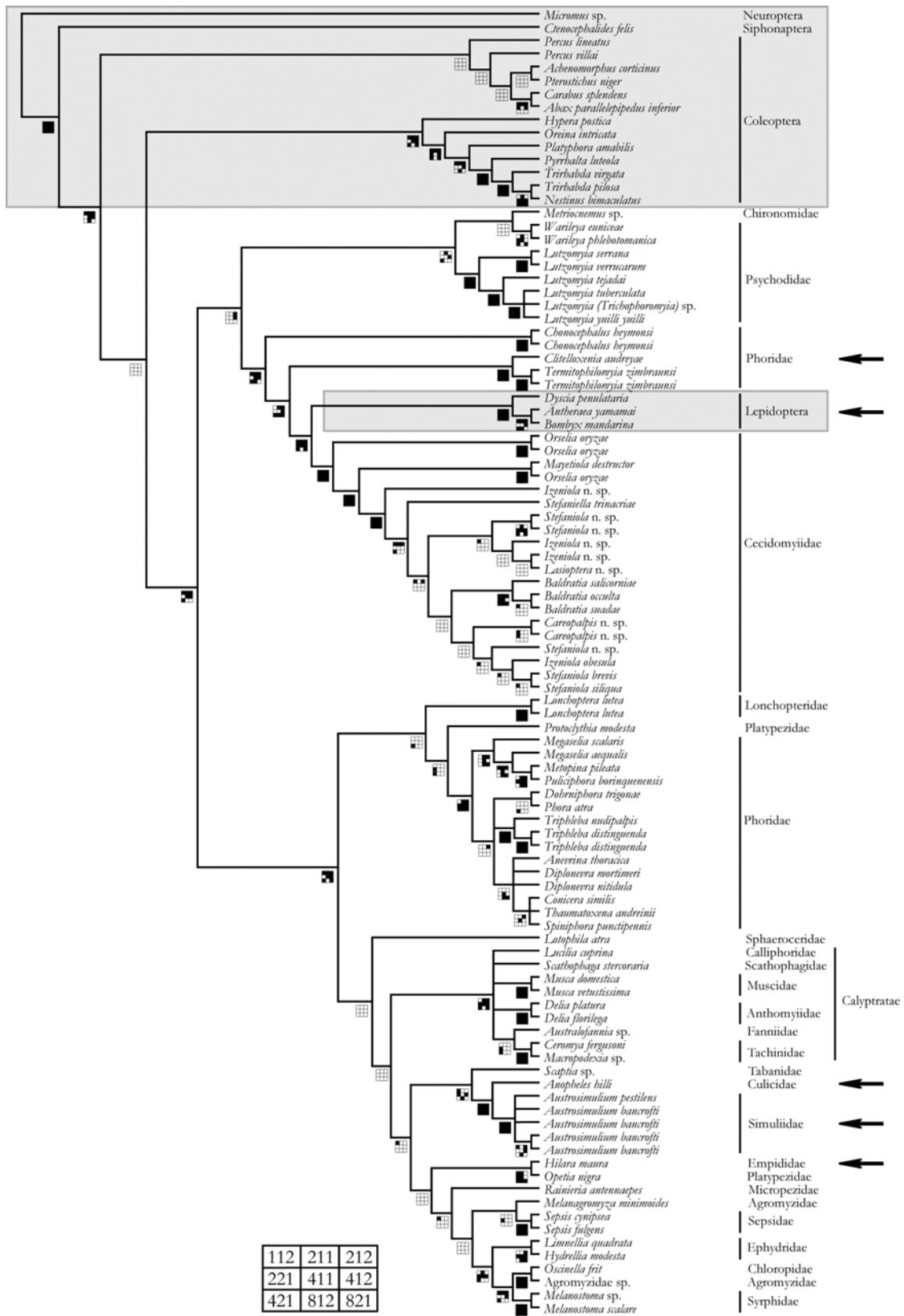
Para o primeiro grupo do gene 12S analisado (Figura 3), em todos os parâmetros de custos utilizados, Diptera resulta como um grupo parafilético. Em virtualmente todos os cladogramas obtidos, Lepidoptera aparece dentro de Diptera como grupo irmão de Cecidomyiidae, que aparece como um grupo monofilético em todos os parâmetros. Esse resultado só não é encontrado com os dados submetidos ao parâmetro de custos 812. Nessa topologia, Psychodidae aparece como grupo irmão de um clado formado por Lepidoptera + (Coleoptera + Diptera). A família Psychodidae resulta como monofilética somente com custos altos para *indels* (parâmetros 812 e 821). Sob todos os outros parâmetros, ela aparece como parafilética, com Chironomidae ora aparecendo como grupo irmão de *Warileya*, ora junto com *Lutzomyia*. Esse último gênero sempre resulta como monofilético, mas o mesmo não acontece com *Warileya*, que não tem sua monofilia sustentada sob os parâmetros que determinam custos maiores para transições (112, 212, 412 e 812). As seqüências dos dois gêneros foram obtidas pelos mesmos pesquisadores (Beati *et al.*, 2004), sendo que, dessa forma, podemos supor que não haja erros de seqüenciamento. Esses autores não incluem em seu trabalho o



posicionamento de *Warileya*, mas os resultados obtidos na presente análise para o gênero *Lutzomyia* não são conflitantes com os resultados obtidos por Beati *et al.* (2004). Em uma análise fenética, Dujardin *et al.* (1999) posicionam *Warileya* na base de Phlebotominae, distante de *Lutzomyia*, que aparece como um dos grupos mais apicais, juntamente com *Phlebotomus*.

Culicomorpha resulta como um grupo parafilético sob todos os parâmetros de custos, com a exclusão de Chironomidae da infraordem. A posição dessa última família varia muito, de acordo com os diferentes custos. Chironomidae aparece como grupo irmão de *Warileya* (Psychodidae) (parâmetro 111), *Lutzomyia* (Psychodidae) (parâmetros 211, 212, 411, 412 e 421), Syrphidae (parâmetro 812) ou *Achenomorphus* (Coleoptera) (parâmetros 221 e 821). O clado Culicidae + Simuliidae está sempre presente, no entanto, duas famílias diferentes se alternam como grupo irmão desse clado, Tabanidae e Sepsidae.

A família Phoridae resulta como parafilética em todas as análises. Os gêneros *Chonocephalus*, *Clitelloxenia* e *Termitophilomyia* sempre aparecem separados do restante da família e muitas vezes formando um clado juntamente com Lepidoptera + Cecidomyiidae. Cook *et al.* (2004) realizaram um estudo dessa família utilizando os genes 12S e 16S. Em sua análise de máxima verossimilhança (*maximum likelihood*), a monofilia de Phoridae é suportada, porém o posicionamento desses três gêneros também é discutido. O comprimento desses ramos é grande e o posicionamento desses três grupos é considerado como sujeito a testes ou como um efeito de *long-branch attraction*. Em nossos resultados, a união desses três gêneros, bem como a sua exclusão de Phoridae parece ser um efeito de *long-branch attraction* (Felsenstein, 1978). As relações entre os gêneros de Phoridae são muito variáveis.



**Figura 3.** Cladograma resultante da análise do primeiro grupo de terminais para o gene 12S, sob o parâmetro 111. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos,

indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

Brachycera aparece como parafilético em todas as análises. Sob o parâmetro 212, esse grupo aparece como monofilético, exceto pela exclusão de parte de Phoridae, que apresentou divergências em todas as análises, como discutido acima (gêneros *Chonocephalus*, *Clitelloxenia* e *Termitophilomyia*). A parafilía desse grupo é definida pela inclusão do clado Culicidae + Simuliidae e pela exclusão de parte de Phoridae.

Ao contrário de Brachycera, Calypttratae é recuperado como um grupo monofilético em quase todos os parâmetros, exceto em dois: 112 e 812. Dentro desse grupo, todas as famílias que apresentaram mais de um terminal na análise resultaram monofiléticas. No entanto as relações entre elas é bastante variável e pouco resolvida.

#### 4.1.2. Segundo grupo de terminais para o gene mitocondrial ribossomal 12S

Os dados das matrizes obtidas com o alinhamento implícito de POY, os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados estão sumarizados na Tabela 4.

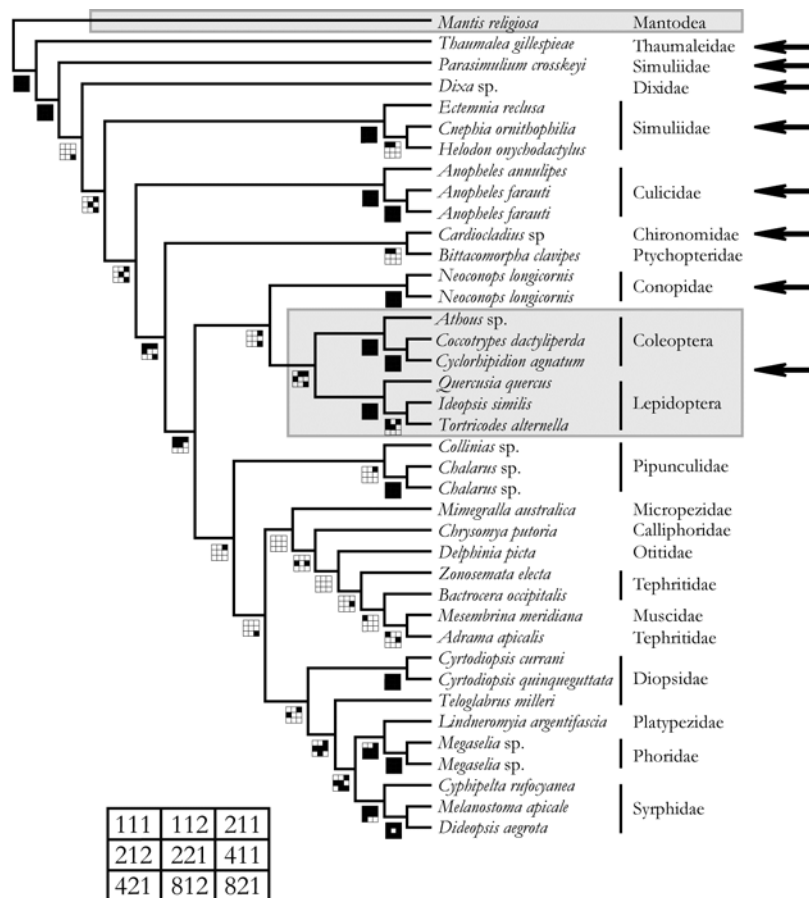
**Tabela 4.** Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o segundo grupo de terminais para o gene 12S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.

Parâmetro	Nº caracteres na matriz	Nº caracteres constantes	Nº caracteres informativos	Nº caracteres não informativos	Comprimento da árvore	Nº de MPT
111	362	89	202	71	949	2
112	411	89	228	94	1168	2
211	334	92	189	53	1054	3
212	340	88	197	55	1323	2
221	354	92	190	72	1574	1
411	322	86	189	47	1212	2
412	326	82	198	46	1491	1
421	330	93	181	56	1777	3
812	317	62	205	50	1740	2
821	319	87	190	42	2068	1

Para esse conjunto de terminais, Diptera nunca resulta como um grupo monofilético (Figura 4). Na maior parte das topologias geradas pelos diferentes parâmetros de custos, os grupos basais de Diptera (Culicomorpha e Ptychopteridae) aparecem na base de um clado formado por Brachycera + Coleoptera + Lepidoptera.

O posicionamento de Coleoptera e Lepidoptera varia muito nos diferentes cladogramas. O clado Coleoptera + Lepidoptera aparece em cinco dos dez parâmetros

analisados (112, 211, 212, 412 e 821), sendo grupo irmão de Conopidae (parâmetros 211, 412 e 821) ou Pipunculidae (112 e 212). Sob outros cinco parâmetros, nos quais Coleoptera e Lepidoptera não formam um grupo monofilético, Lepidoptera aparece como grupo irmão de Conopidae sob quatro deles, sendo que somente no parâmetro 812 essa ordem resulta como grupo irmão de Diptera menos Thaumaleidae. Coleoptera aparece como grupo irmão de Brachycera sob dois parâmetros (111 e 221) e sob outros dois, na base de Lepidoptera e Diptera, menos Thaumaleidae. Thaumaleidae aparece como grupo irmão de Diptera + Lepidoptera + Coleoptera.



**Figura 4.** Cladograma resultante da análise do segundo grupo de terminais para o gene 12S, sob o parâmetro 412. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

Culicomorpha nunca resulta monofilético e Chironomidae aparece como grupo irmão de Ptychopteridae em três parâmetros (111, 112 e 412). A monofilia da família Simuliidae não é

recuperada em nenhuma das topologias. Sob os parâmetros de custos 411, 412, 812 e 821, o gênero *Parasimulium* é excluído dessa família. Em todos os outros parâmetros *Parasimulium* resulta como grupo irmão do restante de Simuliidae. Esse resultado é proposto por Wood & Borkent (1982) para caracteres morfológicos e por Moulton (2000) para dados moleculares. Esse autor afirma que seqüências do gene 28S de *Parasimulium* são alinhadas mais facilmente com seqüências de Thaumaleidae que com outros simulídeos, sugerindo esse fato como uma evidência da distância filogenética de *Parasimulium* dos outros simulídeos, na verdade, da conservação ou plesiomorfia.

Em nenhum dos cladogramas gerados com esse grupo de terminais, Brachycera resulta como um grupo monofilético, seja pela inclusão de Coleoptera e Lepidoptera ou do clado Platypezidae + Culicomorpha, exceto Thaumaleidae.

## **4.2. Gene mitocondrial ribossomal 16S**

Da mesma forma que para o gene 12S, há um grande número de seqüências disponíveis para o gene 16S. Foram então definidos dois grandes grupos para as análises, sendo: (1) um grupo contendo 30 terminais no grupo interno, representando 15 famílias de Diptera, incluindo apenas uma família de dípteros basais (“Nematocera”) e 14 famílias de Brachycera (Anexo III); e (2) um grupo contendo 72 terminais no grupo interno, representando 39 famílias de Diptera, incluindo quatro famílias de dípteros basais e 35 famílias de Brachycera (Anexo IV).

### **4.2.1. Primeiro grupo de terminais para o gene mitocondrial ribossomal 16S**

Os dados das matrizes obtidas com o alinhamento implícito de POY, os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados estão sumarizados na Tabela 5.

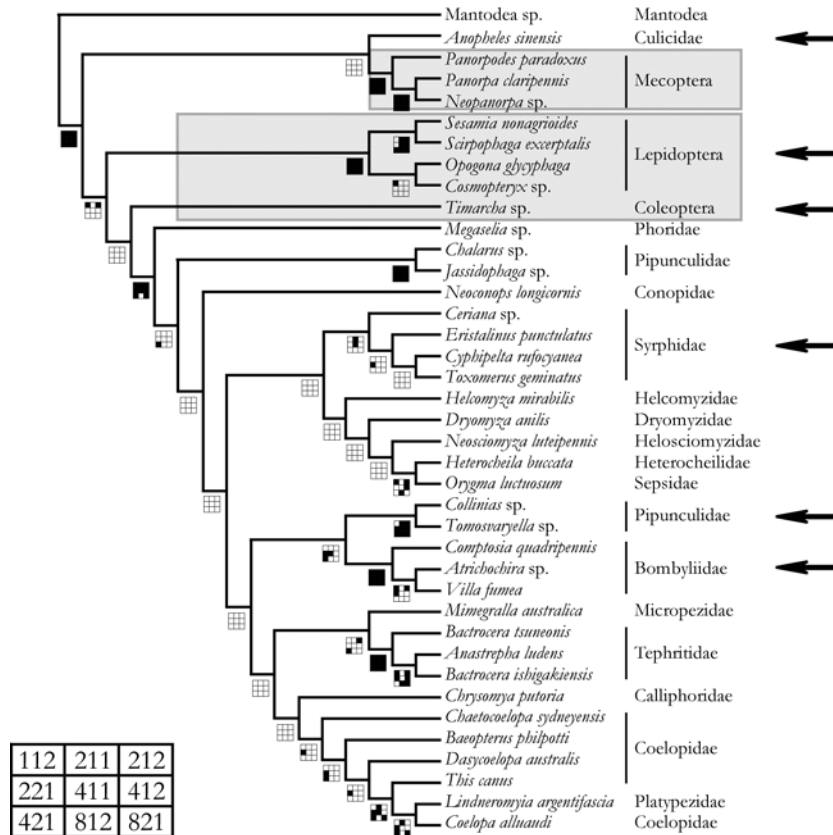
Nesse conjunto de terminais, somente uma espécie de Culicidae foi incluída (Figura 5). Não foi possível incluir nenhum outro terminal de dípteros basais, pois os alinhamentos não eram suficientemente confiáveis. Culicidae aparece excluído de Diptera em todas as topologias, ora como grupo irmão de Mecoptera (parâmetro 111), ora como grupo irmão de Lepidoptera (parâmetros 211 e 421) e, sob vários parâmetros de custos, como grupo irmão de Mecoptera + Lepidoptera + Coleoptera + restante de Diptera (112, 212, 221, 411, 412 e 821). Somente em

uma topologia, essa família aparece mais próxima de Diptera (parâmetro 812), na qual forma um clado com a família Bombyliidae, sendo esse clado grupo irmão de Lepidoptera. Exceto sob esse mesmo parâmetro, Brachycera resulta como um grupo monofilético em todas as outras topologias.

**Tabela 5.** Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o primeiro grupo de terminais para o gene 16S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.

Parâmetro	Nº caracteres na matriz	Nº caracteres constantes	Nº caracteres informativos	Nº caracteres não informativos	Comprimento da árvore	Nº de MPT
111	472	125	227	120	1176	1
112	539	131	255	153	1405	1
211	408	131	201	76	1284	4
212	418	128	209	81	1567	1
221	444	125	219	100	2000	1
411	394	117	203	74	1405	4
412	396	124	206	66	1703	2
421	406	130	204	72	2192	1
812	391	114	213	64	1911	1
821	393	111	209	73	2433	1

As relações entre as famílias de Brachycera variam muito nos diferentes resultados. Não foi recuperada a monofilia de nenhum grande grupo de Brachycera. No entanto, algumas poucas famílias tiveram a sua monofilia mantida, independentemente da variação de custos para os diferentes tipos de mutação. As famílias Bombyliidae e Tephritidae resultaram como monofiléticas em todos os cladogramas, diferentemente de Pipunculidae, que resultou como parafilética em todos os cladogramas. Syrphidae aparece como monofilética em alguns cladogramas (111, 211 e 411), bem como Coelopidae (112, 212, 412 e 812). Aparentemente a monofilia desse último grupo é mantida pelo maior valor de transições com relação às transversões, já que esse resultado aparece em todos os cladogramas gerados por aqueles parâmetros que indicam custo 1 para transversões e custo 2 para transições. Meier & Wiegmann (2002) também relatam que a monofilia de Coelopidae é muito instável com relação à mudança de parâmetros de custos em análises moleculares, embora esse clado seja bem suportado por caracteres morfológicos.



**Figura 5.** Cladograma resultante da análise do primeiro grupo de terminais para o gene 16S, sob o parâmetro 111. Os quadros pintados correspondem aos cladogramas encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos cladogramas. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

A única topologia em que *Eremoneura* (i.e., Empidoidea + Cyclorhapha) aparece como um grupo monofilético é sob o parâmetro 812, no qual Bombyliidae resulta como grupo irmão de Culicidae.

#### 4.2.2. Segundo grupo de terminais para o gene mitocondrial ribossomal 16S

Os dados das matrizes obtidas com o alinhamento implícito de POY, os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados estão sumarizados na Tabela 6.

Esse grupo foi o que incluiu o maior número de famílias de Diptera. No entanto, a diversidade de famílias basais é muito pequena, ficando restrita a apenas quatro grupos, todas elas da infraordem Culicomorpha.



**Tabela 6.** Tamanho da matriz gerada pelo alinhamento implícito de POY, número de caracteres informativos, não informativos e constantes e o comprimento dos cladogramas obtidos para o segundo grupo de terminais para o gene 16S, sob os diferentes parâmetros de custos. MPT = cladogramas igualmente parcimoniosos.

Parâmetro	Nº caracteres na matriz	Nº caracteres constantes	Nº caracteres informativos	Nº caracteres não informativos	Comprimento da árvore	Nº de MPT
111	528	84	261	183	2117	1
112	651	81	295	275	2437	2
211	393	89	211	93	2333	3
212	403	85	218	100	2753	2
221	507	89	246	172	3662	2
411	381	89	205	87	2590	1
412	383	88	207	88	3020	3
421	388	90	205	93	4107	1
812	356	76	221	59	3439	2
821	368	85	210	73	4641	1

Mesmo o grande número de famílias amostradas para esse gene não auxiliou na resolução das relações filogenéticas entre os diferentes grupos, senão apenas incluiu mais ruído na análise (Figura 6).

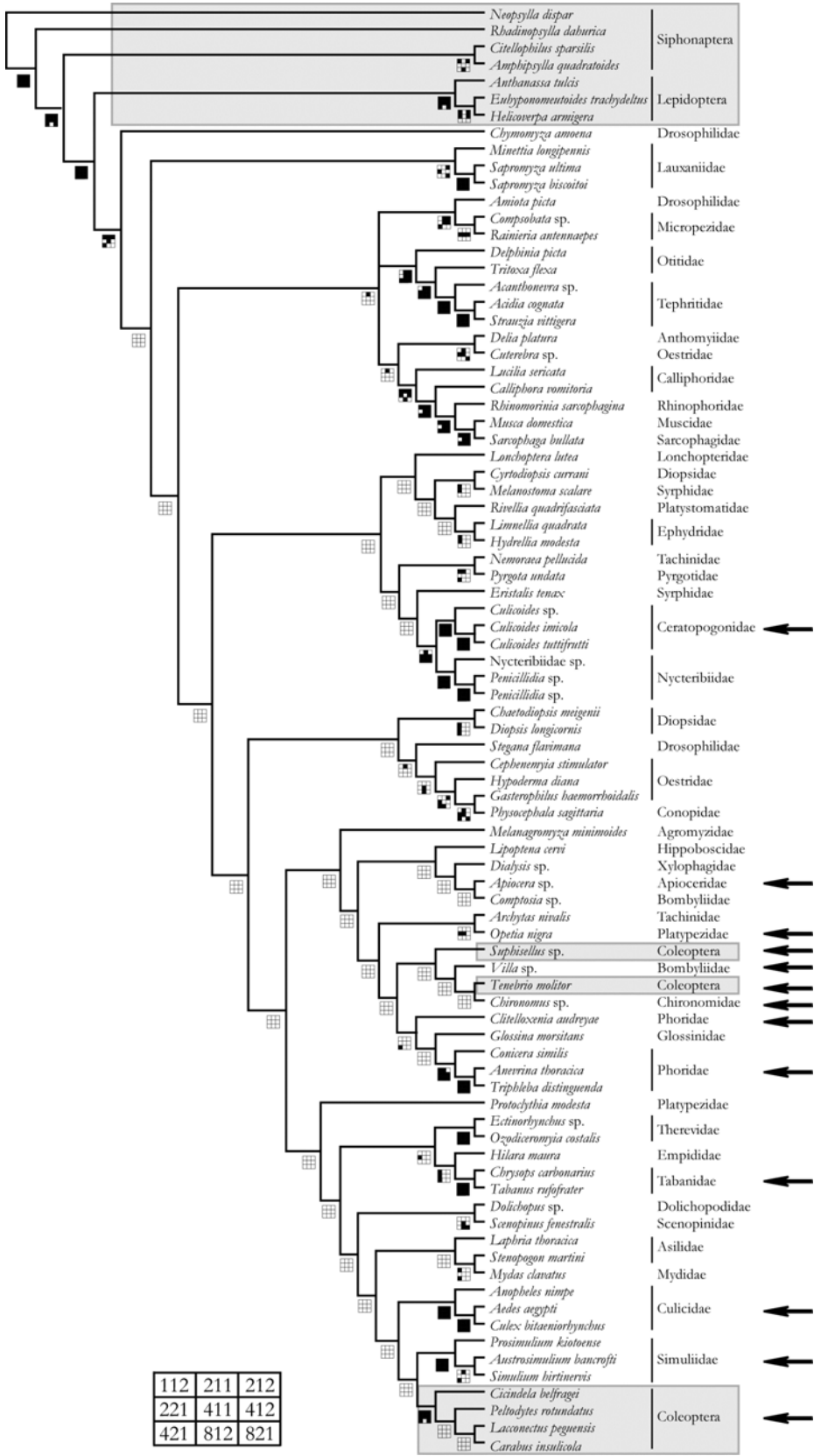
Diptera resulta como parafilético em todos os parâmetros estudados. As espécies de Coleoptera aparecem inseridas em Diptera nas mais diferentes posições na topologia. O clado formado pelos gêneros *Cicindela*, *Peltodytes*, *Lacconectus* e *Carabus* resulta como grupo irmão de Simuliidae (111) ou Culicidae (112 e 212). A família Chironomidae aparece sob alguns parâmetros dentro de Coleoptera (211, 221 e 411).

Absolutamente nenhum dos grandes grupos de Diptera têm suas monofilias recuperadas nessas análises. Os cladogramas gerados são profundamente discordantes com qualquer hipótese morfológica já apresentada ou com outros resultados moleculares, sendo impraticável a descrição das relações filogenéticas dos grupos como apresentadas por essas topologias.

Há uma inversão completa na topologia, com grupos basais de Diptera e grupos externos aparecendo em posições elevadas na topologia e grupos apicais, como famílias de Cyclorrhapha, em posições basais.

Hertwig *et al.* (2004) examinaram seqüências dos genes 12S e 16S em Anura e observaram um alto grau de saturação, mesmo nas regiões mais conservadas desses genes. Esses autores sugeriram que as ambigüidades na homologia primária em grande parte das seqüências e o ruído nos dados foram os responsáveis pelos resultados conflitantes

encontrados e a alta sensibilidade das hipóteses filogenéticas sob os diferentes grupos de parâmetros.



**Figura 6.** Cladograma resultante da análise do segundo grupo de terminais para o gene 16S, sob o parâmetro 111. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

### 4.3. Gene mitocondrial citocromo oxidase subunidade I (COI)

Assim como para os genes mitocondriais 12S e 16S, a grande disponibilidade de seqüências disponíveis para o gene citocromo oxidase subunidade I (COI) permitiu definir dois grupos para as análises: (1) um grupo contendo 43 terminais no grupo interno, representando 27 famílias de Diptera, incluindo nove famílias de dípteros basais (“Nematocera”) e 18 famílias de Brachycera (Anexo V); e (2) um grupo contendo 51 terminais no grupo interno, representando 21 famílias de Diptera, incluindo quatro famílias de dípteros basais e 17 famílias de Brachycera (Anexo VI).

Esses conjuntos de dados foram analisados em partições, isolando-se somente as primeiras, segundas e terceiras posições dos códons. A análise das seqüências completas também foram realizadas e esse particionamento permitiu o cálculo do índice de incongruência.

#### 4.3.1. Primeiro grupo de terminais para o gene mitocondrial COI

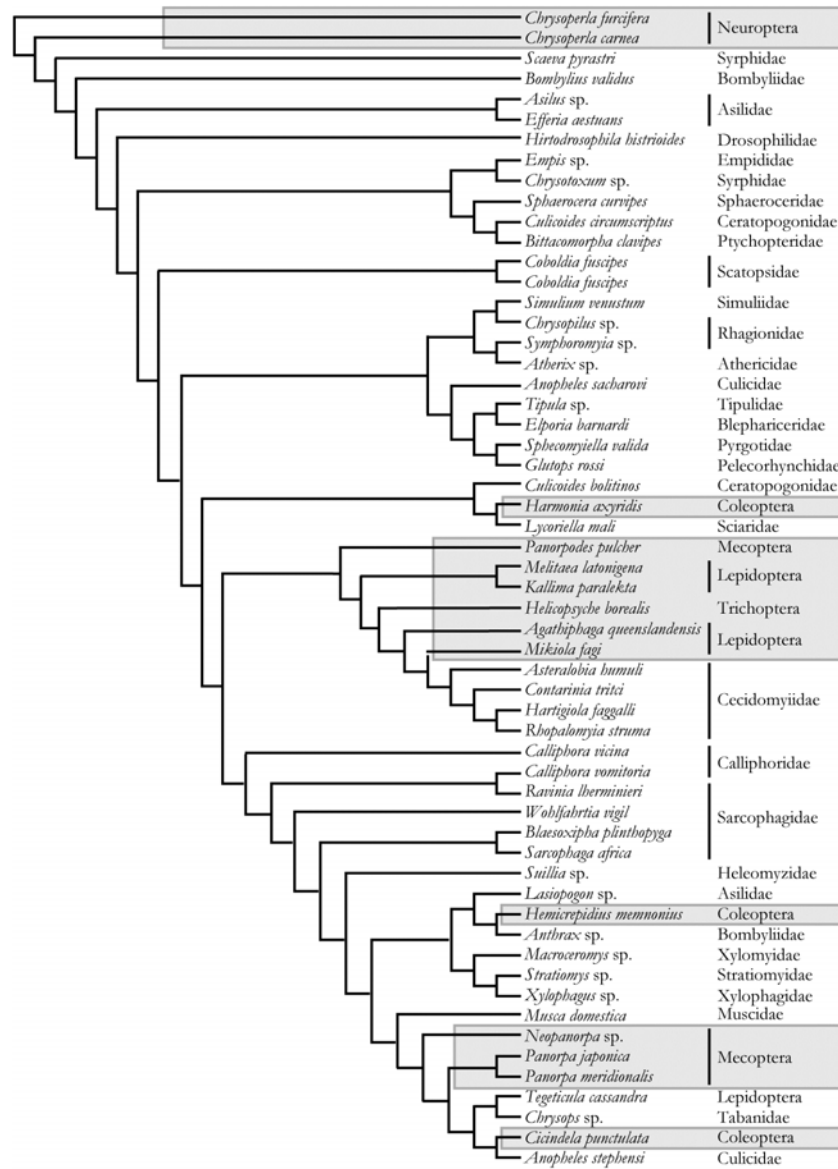
Os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados e o valor de ILD estão sumarizados na Tabela 7.

**Tabela 7.** Comprimento dos cladogramas obtidos para o primeiro grupo de terminais do gene COI, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.

TV:IS	1ª pos.	2ª pos.	3ª pos.	Total	ILD
1:1	343	67	1277	1791	0,058068
2:0	150	54	1662	1976	0,05566802
2:4	1100	160	3182	4786	0,07187631
2:6	1162	180	3196	4874	0,06893722
2:8	1200	184	3200	4932	0,07055961
4:0	300	108	3324	3952	0,05566802
4:2	884	206	4264	5664	<b>0,05473164</b>
4:6	1816	310	5766	8426	0,06337527
6:0	450	162	4986	6006	0,06793207
6:2	1062	274	5940	7700	0,05506494

6:4	1574	342	6832	9274	0,05671771
6:8	2502	444	8326	12036	0,06347624
8:0	600	216	6648	8008	0,06793207
8:2	1230	340	7608	9732	0,05692561
8:6	2260	476	9408	12870	0,05641026

O cladograma apresentado na Figura 7 foi gerado pelo parâmetro 4:2, aquele com o menor valor de ILD. No entanto, todos os valores para esse índice são altos, indicando grande incongruência nas topologias geradas pelas diferentes partições. Verificando a topologia resultante com o parâmetro de custos que gerou o conjunto de dados mais congruente (4:2), observamos que as relações filogenéticas apresentadas são extremamente inconsistentes. Praticamente nenhum clado sabidamente monofilético manteve esse *status*, com exceção de Cecidomyiidae, que resulta como grupo irmão de Lepidoptera. Famílias de Culicomorpha aparecem distribuídas por todo cladograma, relacionadas com outros grupos de dípteros basais, grupos basais de Brachycera (Rhagionidae) ou com táxons pertencentes a outras ordens.



**Figura 7.** Cladograma resultante da análise do primeiro grupo de terminais para o gene COI, sob o parâmetro de custos 4:2. Em cinza, os terminais que não pertencem a Diptera.

Terminais pertencentes aos grupos externos, das ordens Coleoptera, Mecoptera, Lepidoptera e Trichoptera, aparecem dentro de Diptera, relacionados com Bibionomorpha, Asiloidea e Tabanomorpha.

#### 4.3.2. Segundo grupo de terminais para o gene mitocondrial COI

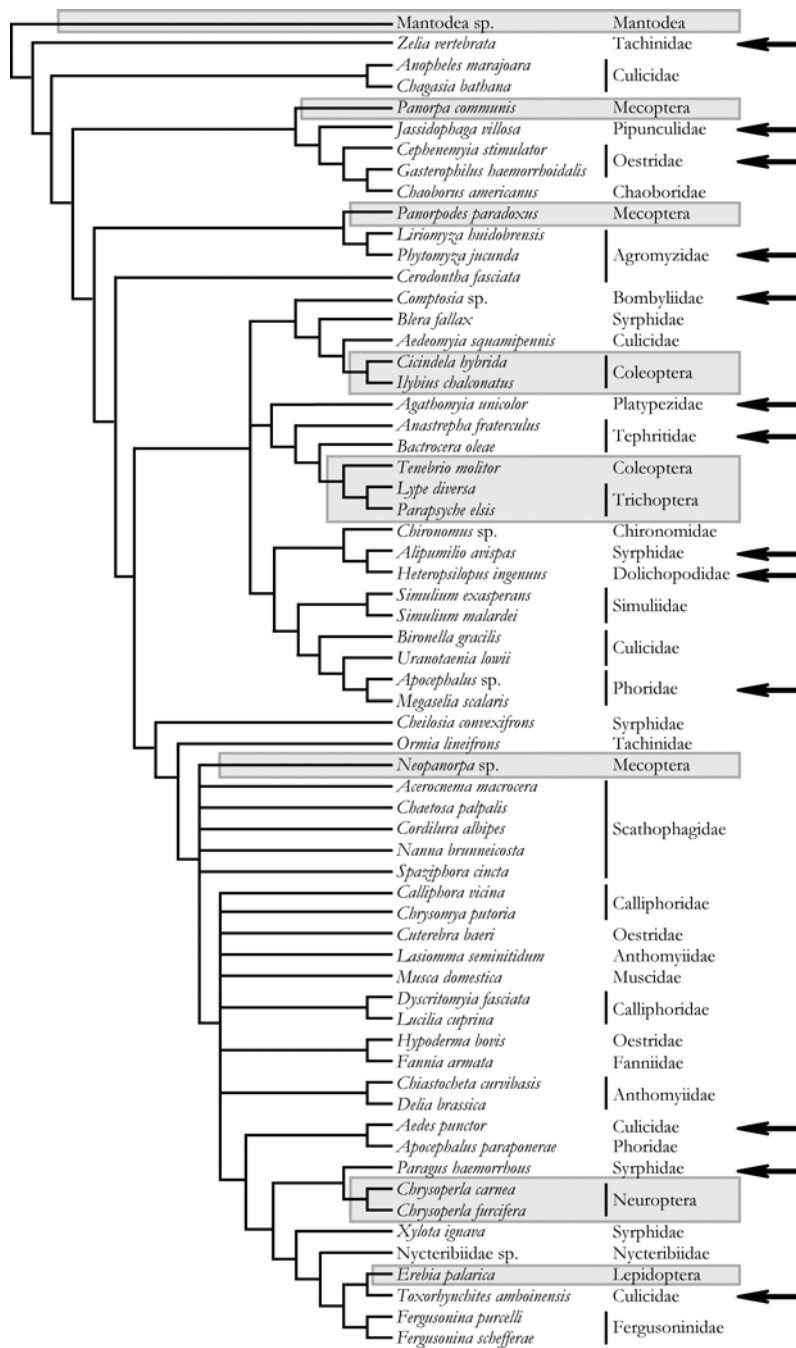
Os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados e o valor de ILD estão sumarizados na Tabela 8.

A exemplo dos resultados obtidos na análise do primeiro grupo de táxons para uma outra região do mesmo gene, o cladograma apresentado na Figura 8 é o gerado pelo parâmetro 2:4, sendo o que apresentou o menor valor de ILD. Os valores desse índice para a presente região do gene COI apresentou valores de ILD altos, indicando grande incongruência entre as partições.

**Tabela 8.** Comprimento dos cladogramas obtidos para o segundo grupo de terminais do gene COI, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.

TV:TS	1ª pos.	2ª pos.	3ª pos.	Total	ILD
1:1	35	929	213	1284	0,083333
2:0	34	1138	112	1414	0,09193777
2:4	78	2418	684	3458	<b>0,08039329</b>
2:6	86	2424	754	3596	0,09232481
2:8	86	2440	812	3690	0,09539295
4:0	68	2276	224	2828	0,09193777
4:2	114	3036	558	4052	0,08489635
4:6	156	4306	1126	6092	0,08273145
6:0	54	60	72	4242	0,95615276
6:2	156	4172	678	5502	0,09014904
6:4	184	4908	990	6628	0,08237779
6:8	226	6170	1554	8662	0,08219811
8:0	72	80	96	5656	0,95615276
8:2	198	5330	794	6934	0,08826074
8:6	254	6776	1418	9194	0,08113987

A topologia mais congruente apresenta relações filogenéticas inconsistentes com todo conhecimento entomológico disponível. Novamente, terminais pertencentes aos grupos externos (Coleoptera, Mecoptera, Lepidoptera, Trichoptera e Neuroptera) aparecem espalhados dentro de Diptera. Apenas três grupos incluídos na análise tiveram a monofilia recuperada: Trichoptera, Neuroptera e a família Fergusoninidae, de Acalyptratae (Diptera). No entanto, o posicionamento desses grupos na filogenia é inconsistente. Fergusoninidae, família incluída em um dos grupos mais apicais de Diptera, resulta como grupo irmão de um clado formado por Lepidoptera + Culicidae.



**Figura 8.** Cladograma resultante da análise do segundo grupo de terminais para o gene COI, sob o parâmetro de custos 2:4. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

#### 4.4. Gene mitocondrial citocromo oxidase subunidade II (COII)



Da mesma forma como foi feito com os grupos analisados do gene COI, os dados do gene COII foram analisado em partições, isolando-se somente as primeiras, segundas e terceiras posições dos códons e também foi procedida a análise das seqüências completas. Esse particionamento permitiu o cálculo do índice de incongruência (ILD). Os terminais utilizados nessa análise estão listados no Anexo VII.

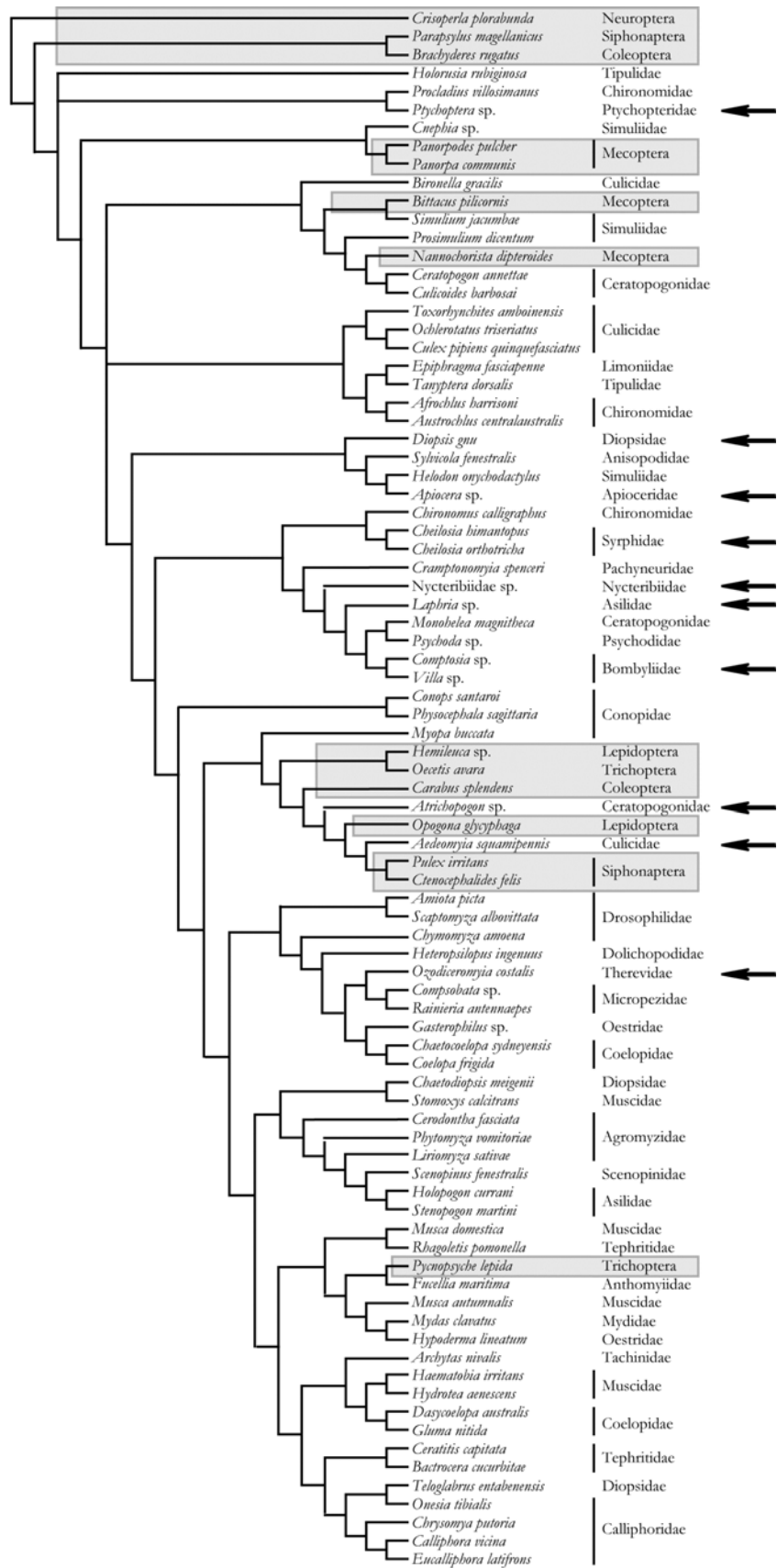
Os comprimentos dos cladogramas obtidos, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados e o valor de ILD estão sumarizados na Tabela 9.

**Tabela 9.** Comprimento dos cladogramas obtidos para o gene COII, das partições e da seqüência completa, sob os diferentes parâmetros de custos. As partições correspondem à posição do códon ativado em cada análise. Negrito indica o menor valor da ILD para o conjunto de dados.

TV:TS	1ª pos.	2ª pos.	3ª pos.	Total	ILD
1:1	1401	447	135	2178	0,089532
2:0	0	18	48	2244	0,97058824
2:4	3618	1344	298	5860	0,10238908
2:6	3648	1458	354	6112	0,10667539
2:8	3688	1536	358	6238	0,10516191
4:0	0	36	96	4488	0,97058824
4:2	4502	1190	414	6742	0,09433403
4:6	6478	2274	620	10414	0,10005761
6:0	4866	738	372	6732	0,11229947
6:2	6172	1464	550	9034	0,09386761
6:4	7354	2092	684	11116	<b>0,08870097</b>
6:8	9284	3180	890	14804	0,0979465

Da mesma forma que nas outras análises de genes mitocondriais, Diptera aparece como parafilético na topologia encontrada pelo parâmetro de custos mais congruente (6:4). Esse conjunto de dados foi o que apresentou, em geral, os maiores valores para ILD, indicando pouca congruência entre as partições de dados.

Novamente a topologia apresenta resultados que podem ser considerados absurdos (Figura 9). Lepidoptera, Mecoptera, Siphonaptera, Coleoptera e Trichoptera aparecem como grupos irmãos de diferentes famílias de Culicomorpha, distribuídos por praticamente todo o cladograma. As únicas famílias que têm sua monofilia recuperada são Syrphidae, Bombyliidae e Micropezidae, o que é irrelevante frente ao restante dos resultados. Nenhum grande grupo aparece como monofilético e as relações entre eles são totalmente discordantes de qualquer outra filogenia já proposta.



**Figura 9.** Cladograma resultante da análise do gene COII, sob o parâmetro de custos 2:4. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

#### **4.5. Robustez das topologias com dados dos genes mitocondriais 12S, 16S, COI e COII**

Esses quatro genes foram os que proporcionaram os conjuntos de dados com maior diversidade de famílias de Diptera para as análises, mas há ainda uma amostragem pequena de grupos basais de Diptera. Aparentemente, uma maior amostragem dessas famílias não resolveria os problemas encontrados nas reconstruções filogenéticas, já que várias ordens utilizadas como grupos externos acabaram por serem incluídas em Diptera. Nesse caso, os genes utilizados é que não fornecem um sinal filogenético compatível com os níveis da filogenia consideradas neste estudo.

Seqüências do gene 16S têm sido estudadas para a reconstrução filogenética de alguns grupos de insetos (e.g., Kambhampati & Charlton, 1999; Kambhampati *et al.*, 2000; Meier & Wiegmann, 2002; Damgaard, 2005). Gimeno *et al.* (1997) compararam hipóteses elaboradas pelos genes 28S e 16S para Braconidae (Hymenoptera) e encontraram um maior suporte dos clados fornecido pelo primeiro gene. No entanto, esses autores também levantam a hipótese de subamostragem em seus dados. Dados semelhantes foram encontrados para Diopsidae (Diptera) (Baker *et al.*, 2001), sendo que os genes nucleares apresentaram maior resolução, menos homoplasia e maior suporte. Esses autores destacam principalmente a incongruência dos genes COII e 12S com o restante dos genes utilizados. Uma comparação do efeito dos genes ribossomais 18S e 12S na filogenia do grupo de espécies de *Anopheles punctulatus* (Beebe *et al.*, 2000) demonstrou que o gene nuclear é mais informativo que o mitocondrial.

Da mesma forma que o observado para anuros (Hertwig *et al.*, 2004), o poder de resolução limitado dos genes 12S e 16S para a filogenia de Diptera pode ser uma consequência das altas taxas de evolução ou diferentes taxas de substituição entre táxons (Simon *et al.* 1994).

#### **4.6. Análise dos genes 18S e 28S**

Os táxons terminais utilizados nessa análise estão listados nos Anexos VIII e IX. Os comprimentos dos cladogramas obtidos nessas análises, bem como o número de cladogramas igualmente parcimoniosos para cada parâmetro de custos utilizados e o valor de ILLD estão sumarizados na Tabela 10.

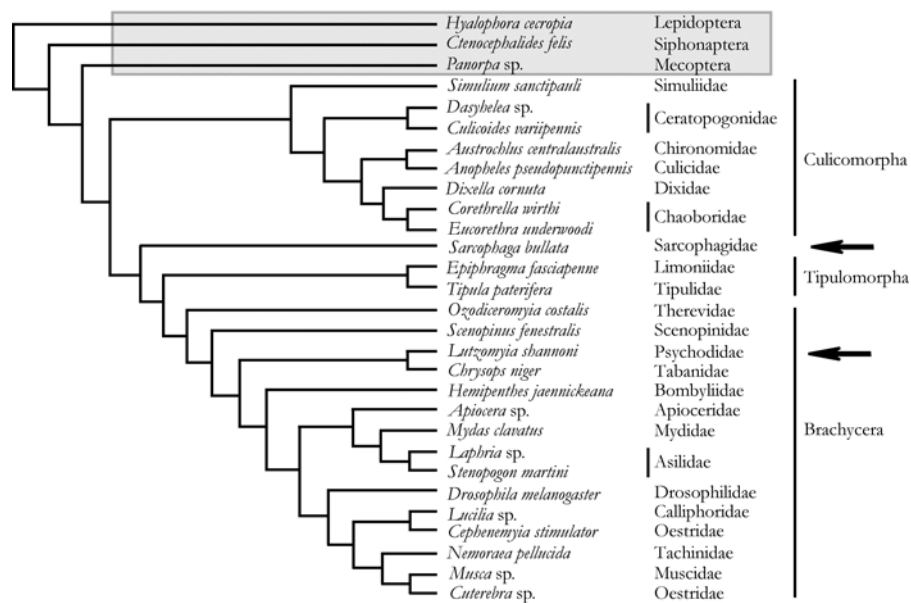
**Tabela 10.** Comprimento dos cladogramas obtidos para os genes 18S e 28S, separadamente e em análise combinada, sob os diferentes parâmetros de custos. Os números entre parênteses indicam o número de cladogramas igualmente parcimoniosos encontrados. Negrito indica o menor valor da ILLD para o conjunto de dados.

Parâmetro	18S	28S	18S + 28S	ILLD
111	1193 (1)	586 (6)	1809 (1)	<b>0,01658</b>
112	1413 (1)	753 (3)	2206 (3)	0,01813
211	1529 (2)	691 (7)	2322 (1)	0,04392
212	1864 (1)	924 (3)	2905 (1)	0,04027
221	1918 (2)	868 (4)	2870 (1)	0,02926
411	1932 (2)	861 (12)	3024 (1)	0,07638
412	2326 (2)	1105 (2)	3676 (3)	0,06664
421	2473 (1)	1073 (1)	3735 (3)	0,05060
812	2765 (2)	1523 (2)	4652 (1)	0,07824
821	2966 (1)	1325 (1)	4788 (1)	0,10380

Dos dez parâmetros de custos analisados, o que apresentou o menor valor de ILLD foi o parâmetro 111 (ILLD = 0,0165), isto é, foi o parâmetro que apresentou a maior congruência entre os dois conjuntos de dados. Isso significa que a ausência de custos diferenciados para os diferentes tipos de mutação corresponde à análise que resulta na melhor congruência entre as análises separadas. Esse mesmo parâmetro de custos foi encontrado por outros pesquisadores como o mais congruente (e.g., Giribet *et al.*, 2000; Meier & Wiegmann, 2002, Damgaard *et al.*, 2005). As topologias resultantes das análises sob esse parâmetro de custos para os genes 18S, 28S e para a análise combinada desses dois genes são apresentadas e discutidas.

#### 4.6.1. Análise do gene 18S

A análise do gene 18S (Figura 10) discorda em alguns pontos bastante relevantes das análises realizadas até agora com dados morfológicos, no entanto, é a primeira análise na qual a ordem Diptera resulta como um grupo monofilético. Culicomorpha e Tipulomorpha resultam como grupos monofiléticos, em acordo com as hipóteses geradas com base em caracteres morfológicos. No entanto, as relações entre as famílias de Culicomorpha são conflitantes com as hipóteses já existentes, nas quais são consideradas duas superfamílias: Culicoidea e Chironomoidea (Hennig, 1981; Wood & Borkent, 1989). Essas superfamílias não são recuperadas nessa análise.



**Figura 10.** Cladograma resultante da análise do gene 18S, sob o parâmetro de custos 111.

Mecoptera resulta como grupo irmão de Diptera, concordando com a hipótese de Hennig (1981), embora o sinal filogenético disponível, por causa da baixa amostragem de grupos externos, seja fraco para uma hipótese mais robusta.

Brachycera, grupo bem sustentado nas análises com dados morfológicos, aparece como parafilético. Sarcophagidae aparece junto com os dípteros basais e Psychodidae como grupo irmão de Tabanidae. Considerando todo o conhecimento de morfologia, fisiologia, embriologia e ecologia de dípteros, esse é uma hipótese muito pouco verossímil.

#### 4.6.2. Análise do gene 28S

Novamente Diptera resulta como um grupo monofilético. O cladograma resultante para as seqüências do gene 28S apresenta Tipulomorpha, Culicomorpha e Brachycera como grupos monofiléticos. No entanto, as relações filogenéticas entre os grandes grupos, bem como dentro deles mesmos não são bem resolvidas (Figura 11). Drosophilidae aparece dentro de Calyptratae, como grupo irmão de Oestridae + Sarcophagidae. Psychodidae aparece como grupo irmão de todos os outros grupos de Diptera, uma hipótese ausente em qualquer outra filogenia proposta até agora.

Para esse gene, o grupo irmão de Diptera é o clado Mecoptera + Siphonaptera, como hipotetizado por Kristensen (1975).

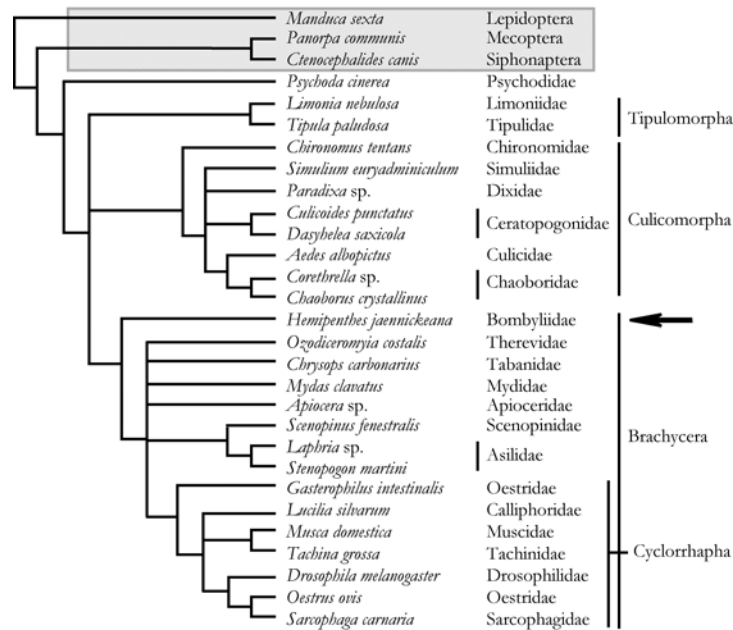


Figura 11. Árvore de consenso gerada pela análise do gene 28S, sob o parâmetro de custos 111.

#### 4.6.3. Análise combinada dos genes 18S e 28S

A análise combinada desses dois conjuntos de dados (genes ribossomais 18S e 28S) resultou em um cladograma no qual Brachycera é parafilético, pois Psychodidae (Psychodomorpha) aparece como grupo irmão de Tabanidae (Brachycera, Tabanomorpha) (Figura 12).

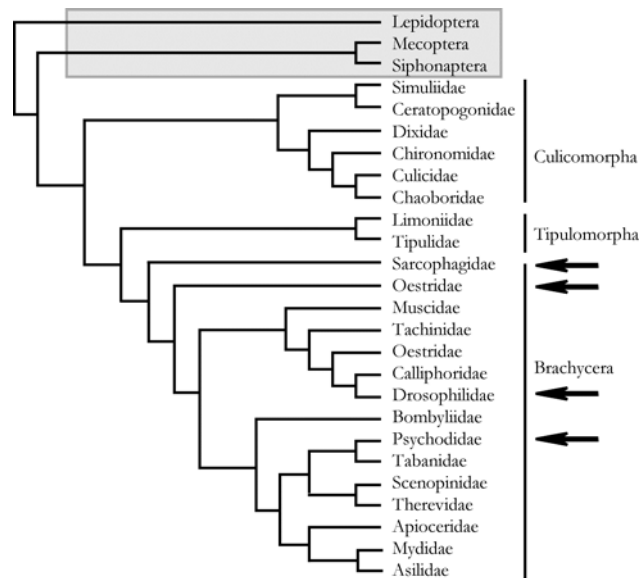


Figura 12. Cladograma resultante da análise combinada dos genes 18S e 28S, sob o parâmetro de custos 111.

Sarcophagidae não aparece mais junto com os Diptera basais, mas na base de Brachycera. A relação entre Psychodomorpha e Brachycera, sendo Tabanidae um dos grupos basais desse grupo, encontra sustentação em algumas características morfológicas de adultos e larvas (Sinclair, 1992; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996). Entretanto, essa posição para Sarcophagidae é equivocada, já que essa família está incluída em Calyptratae, um dos grupos mais apicais na filogenia de Diptera (McAlpine, 1989; Oosterbroek & Courtney, 1995; Yeates & Wiegmann, 1999).

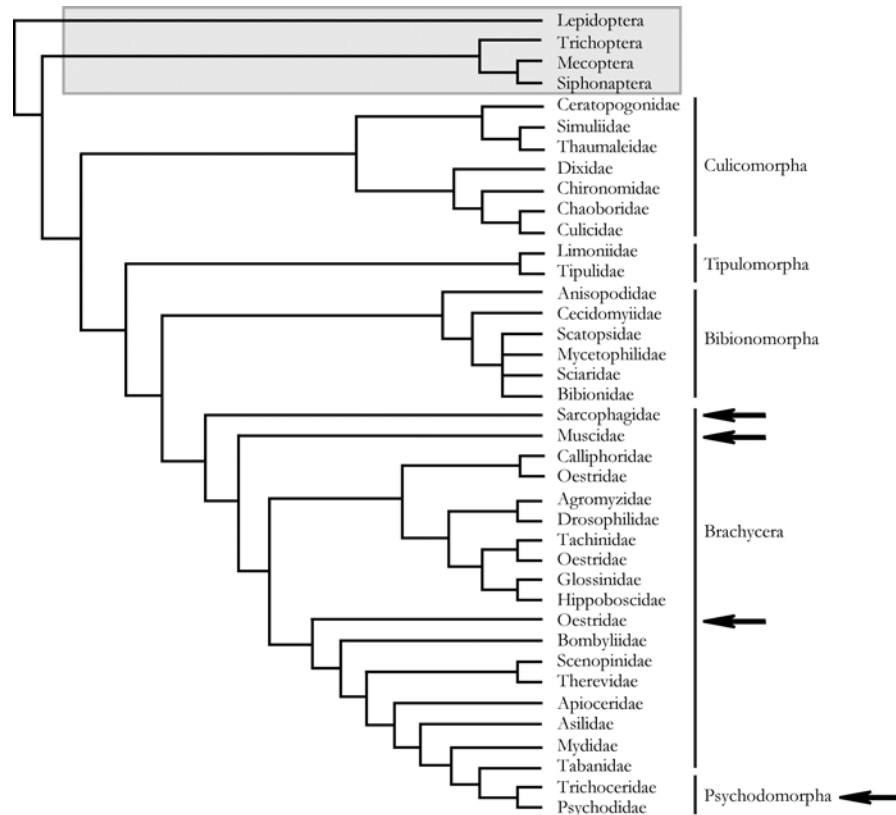
#### **4.6.4. Análise combinada dos genes 18S e 28S com o aumento no número de táxons terminais, com inclusão de *missing data***

Com o aumento do número de terminais na análise, e com a conseqüente inserção de *missing data* (Figura 13), alguns grupos basais de Diptera, como Culicomorpha e Tipulomorpha aparecem monofiléticos. Scatopsidae é incluído em Bibionomorpha, recuperando a hipótese de Hennig (1981). No entanto, esse posicionamento de Scatopsidae é atualmente considerado improvável, sendo essa família inserida na infraordem Psychodomorpha (Wood & Borkent, 1989; Courtney, 1990, 1991; Sinclair, 1992; Amorim, 1994; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996; Amorim & Grimaldi, 2006).

As distorções em Brachycera são bastante profundas. Não apenas o grupo é parafilético em relação aos Psychodomorpha, mas este grupo está em uma posição terminal na filogenia. De fato, há uma inversão completa na filogenia interna de Brachycera, com as famílias de Calyptratae formando um grupo basal e os grupos sucessivamente mais basais em uma posição mais apical. Esse resultado pode ter sido influenciado pelas seqüências do primeiro gene (18S), já que a análise desse único gene também apresenta Brachycera como um grupo parafilético. Como a inclusão de *missing data* não afeta esse ponto da filogenia, a taxa de evolução do gene, bem como problemas metodológicos, como o alinhamento das seqüências, podem ser alguns dos fatores que levam a esses resultados.

É interessante observar que, a despeito da inserção de grande quantidade de *missing data* (21 dos 39 terminais não tinham uma das duas seqüências, em 12 faltando a seqüência do gene 18S e em 9, do gene 28S), os resultados gerados não foram muito diferentes das outras análises no que diz respeito à monofilia dos grandes grupos. Nenhuma das famílias de Bibionomorpha incluídas na análise apresentava dados para o gene 18S, o que faz com que as

relações internas entre as famílias que o compõem na análise conjunta sejam ditadas exclusivamente pelos resultados com o gene 28S.



**Figura 13.** Cladograma resultante da análise combinada dos genes 18S e 28S, sob o parâmetro de custos 111, com a inclusão de *missing data*.

#### 4.7. Análise do gene 28S com a inclusão da ordem Strepsiptera

Existe na literatura a sugestão de que o grupo irmão de Diptera seja a ordem Strepsiptera (Whiting & Wheeler, 1994; Whiting *et al.*, 1997). Para testar essa hipótese, incluímos na matriz de dados para o gene 28S uma seqüência de *Stylops mellitae* (Figura 14). O conjunto das topologias obtidas mostra grande variação para as relações internas em Diptera, bem quanto a hipóteses de monofilia de grandes grupos, inclusão ou exclusão de famílias particulares nas subordens; e relações filogenéticas entre famílias dentro das subordens.

Mecoptera sempre aparece como monofilético, enquanto que Siphonaptera aparece como parafilético sob o parâmetro de custos 821. Trichoptera aparece em duas posições, como grupo irmão de Mecoptera + Siphonaptera ou como grupo irmão de Mecoptera +



Siphonaptera + Strepsiptera + Diptera. Mecoptera aparece ora como grupo-irmão de Siphonaptera, ora como um clado independente, ora com Siphonaptera e Trichoptera.

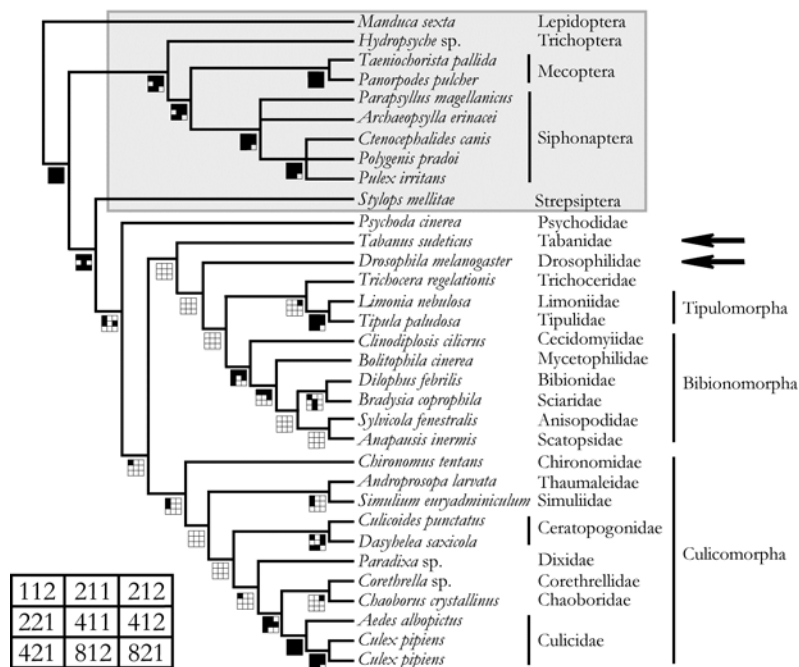
Strepsiptera corresponde a um caso especial. Entre os autores usando dados morfológicos, a posição desse grupo na filogenia de Holometabola não era objeto de consenso, aparecendo ora como grupo-irmão de Hymenoptera, ora próximo a Coleoptera (Whiting, 1998). Recentemente surgiram hipóteses nas quais Strepsiptera aparece como grupo-irmão de Diptera com base em dados moleculares (Whiting *et al.*, 1997; Whiting, 2002). Nas análises realizadas no presente trabalho, mesmo com a introdução de grupos externos a Holometabola, Strepsiptera aparece como grupo-irmão de Diptera ou - na maioria dos casos - como um subgrupo de Culicomorpha.

O estudo cuidadoso da seqüência de 28S de Strepsiptera utilizada (a única disponível até o momento que alinhou com as seqüências utilizadas) permite compreender melhor a causa desse padrão: o alinhamento de grande parte do trecho utilizado é virtualmente incompatível com todas as demais seqüências – efeito de *long-branch attraction*. Isso implica que o número de passos para passar da seqüência de qualquer espécie terminal (ou grupo monofilético mais inclusivo) para a seqüência de Strepsiptera é tão grande que a determinação de sua posição na topologia se torna extremamente difícil e arbitrária. Há uma de duas explicações possíveis para esse quadro: (1) erro na determinação da seqüência; ou (2) a evolução do grupo foi de tal modo drástica que a comparação entre as seqüências e, conseqüentemente, as hipóteses de homologia entre as bases têm pouco valor. Nenhum desses casos fornece corroboração para a hipótese de Strepsiptera como grupo-irmão de Diptera. Outros autores, mais recentemente, também discordam dessa hipótese de relação entre Diptera e Strepsiptera em análises com dados moleculares (Hwang *et al.*, 1998; Rokas *et al.* 1999; Hayward *et al.*, 2005). Bridge *et al.* (2003) sugerem que cerca de 20% das seqüências disponíveis para fungos não são confiáveis, ou seja, foram publicadas com erros (erro na identificação do material, seqüências dúbias ou quiméricas). Dessa forma, pode existir a possibilidade de essa seqüência de Strepsiptera ser erroneamente determinada.

O grupo mais basal de Diptera varia muito entre os diferentes parâmetros de custos. Culicomorpha, grupo considerado muito sólido na literatura, aparece como monofilético somente sob três parâmetros de custos (111, 112 e 221). Brachycera aparece como monofilético na maior parte das análises (211, 221, 411, 412, 421, 812 e 821). Bibionomorpha (incluindo Scatopsidae) aparece como monofilético na maior parte dos casos, mas sob alguns

parâmetros esse clado resulta como parafilético (411, 812 e 821). Psychodomorpha usualmente aparece como parafilético, uma vez que Scatopsidae aparece em Bibionomorpha em grande número de casos (exceto sob os parâmetros 411, 812 e 821) e Trichoceridae aparece alternativamente em Psychodomorpha (221, 411, 412 e 421) e em Tipulomorpha (111 e 212).

O cladograma apresentado na Figura 11 é resultado de análise com pesos iguais para transversões, transições e *indels*. Nesse cladograma, Brachycera e Psychodomorpha aparecem como parafiléticos e as relações entre as subordens é inverossímil considerando o conhecimento acumulado de morfologia. O uso de pesos iguais para as diferentes mudanças, resulta em grande inconsistência em relação ao conhecimento de filogenia de Diptera com base em morfologia.



**Figura 14.** Cladograma resultante da análise do gene 28S, sob o parâmetro de custos 111, com a inclusão de Strepsiptera no grupo externo. Os quadros pintados correspondem aos clados encontrados nas análises com os diferentes parâmetros de custos, indicados nos quadros numerados. Os quadros em branco indicam a ausência dos respectivos clados. Em cinza, os terminais que não pertencem a Diptera. As setas indicam os grupos cuja posição foge muito de outras reconstruções.

#### 4.8. Robustez das topologias com dados dos genes nucleares 18S e 28S

Os genes 18S e 28S foram estudados por Friedrich & Tautz (1997b) e são citados por esses autores como bons indicadores da filogenia de Diptera. Ainda assim, esses mesmos

autores sugerem evidências de que as taxas evolutivas desses dois genes são maiores em Díptera que em outros grupos de Holometabola.

Aparentemente, as limitações de subamostragem de táxons internos e externos podem ser responsabilizadas por parte das diferenças dos resultados nas análises aqui realizadas com esses dois genes em relação às topologias propostas a partir de caracteres morfológicos. A ausência de uma amostragem maior de espécies e de níveis de táxons externos lança suspeição sobre toda a otimização de caracteres (tanto no alinhamento, quanto na determinação das relações filogenéticas) na base de Díptera. Isto é, o uso de apenas três níveis externos compromete a confiança na determinação de qual seja o grupo-irmão do restante dos Díptera, o que, por sua vez, pode comprometer o sinal na determinação da primeira heterobatmia em Díptera. Por outro lado, a subamostragem de táxons internos, ou seja, o número pequeno de espécies amostradas em cada família (na maior parte dos casos, uma única espécie), tende a provocar o efeito de *long-branch attraction*, indevidamente transferindo para a base do clado da família características que, de fato, pertencem a níveis mais recentes na evolução do grupo. Adicionalmente, o uso de mais de uma espécie de cada uma das (enormes) famílias faria com que homoplasias entre espécies fossem diluídas, de maneira a evitar que as homoplasias, através do processo de otimização, alcançassem clados inferiores, desse modo diminuindo o ruído na determinação das relações entre clados maiores.

Os resultados alcançados neste trabalho mostram, portanto, que análises conjuntas de genes poderiam reduzir parte dos desvios que resultam do uso de uma única fonte de dados (isto é, de um único gene) no estabelecimento de relações de parentesco. No entanto, a ausência de dados para todos os clados principais envolvidos (*missing data*) e a subamostragem de clados internos e externos pode acabar resultando em desvios no processo de otimização, gerando alinhamentos deficientes e, ao final, a falsa associação entre clados. Nem mesmo o processo de análise conjunta de genes pode corrigir esses problemas. Baseado em grupos de dados simulados, Wiens (1998) demonstrou que a adição ou combinação de grupos de caracteres incompletos geralmente aumenta ou tem pouco efeito no decréscimo da resolução das reconstruções filogenéticas. Talvez a associação de um número grande de genes possa reduzir em um grau algumas dessas deficiências, mas corrigir as limitações de subamostragem parece ter um efeito mais forte na qualidade dos resultados finais que propriamente a associação de vários conjuntos deficientes de dados.



	1	380	115	221	56-77	5	626	0,6438	0,7015
	2	361	139	189	37-58	9	590	0,6085	0,6984
		354	135	187	30-51	8	590	0,6051	0,7013
		354	135	187	30-51	4	590	0,6051	0,7020
	3	389	133	191	65-86	9	539	0,7069	0,7676
<b>F</b>									
	1	708	204	391	117-170	1	1398	0,5708	0,5848
	2	653	233	353	62-115	1	1378	0,5443	0,5977
		646	229	351	55-108	1	1369	0,5464	0,6057
		646	229	351	55-108	1	1368	0,5468	0,6068
	3	707	227	347	116-169	1	1202	0,6656	0,6905

Em todas as análises, os alinhamentos com menor número de *indels* foram obtidos pelo programa MALIGN. CLUSTAL gerou os alinhamentos com tamanhos intermediários, enquanto que o POY gerou os alinhamentos com maior número de caracteres (portanto, *indels*) e menor número de passos dos cladogramas finais.

Como foi comentado, o programa CLUSTAL calcula uma árvore de distância para as seqüências em análise construindo uma árvore guia, buscando os terminais na seqüência em que os dados foram colocados na matriz. Como a construção do alinhamento é feita com base nessa árvore, não há uma busca maior por alinhamentos mais parcimoniosos. Dessa forma, o número de caracteres constantes na matriz gerada é o menor entre os três programas. Isso demonstra que o número de mudanças de bases não é minimizado pelo programa. No entanto, não foi o programa que inseriu o maior número de *indels* nos alinhamentos.

Os alinhamentos gerados pelo POY foram os que tiveram o maior número de *indels* inseridos. Mesmo sendo a maior matriz gerada pelos programas aqui utilizados, foram os alinhamentos que geraram os cladogramas mais parcimoniosos. Os números de caracteres constantes e informativos nas matrizes são muito semelhantes aos encontrados pelo MALIGN. No entanto, os cladogramas gerados por esse último programa possuem maior número de passos que os encontrados a partir das matrizes geradas pelo POY. Os índices de consistência e retenção encontrados para os cladogramas resultantes dos alinhamentos gerados pelo MALIGN são os menores para os três programas em praticamente todas as análises.

Ao compararmos um trecho da seqüência da região D1 do gene 28S alinhado pelos programas MALIGN e POY, verificamos que esse último programa inseriu um número muito maior de *indels* que o primeiro (Figura 15). Com a inserção de um grande número de *indels*, o programa POY conferiu à matriz um número muito maior de caracteres autapomórficos, que não são informativos em uma análise filogenética (24 para POY, contra 5 para MALIGN, somente para esse trecho do alinhamento). Para esse mesmo trecho, no entanto, os dois alinhamentos possuem um número semelhante de caracteres informativos, 39 para MALIGN

e 35 para POY - e, no alinhamento completo, esse número é muito próximo (163 caracteres informativos na matriz gerada por POY e 164, por MALIGN).

Os cladogramas resultantes das análises dessas duas matrizes têm comprimentos muito distintos: 635 passos para POY, contra 710 passos para MALIGN. Com um maior número de *indels* inseridos, ocorre uma diminuição no número de estados encontrados em um caráter. Por exemplo, nesse pequeno trecho (Figura 15) encontramos dois caracteres (44 e 60) que possuem os cinco estados de caráter (A, C, G, T e *indel*) no alinhamento de POY (marcados em vermelho), enquanto que, na matriz gerada por MALIGN, encontramos quatro caracteres que possuem os cinco estados (28, 29, 35 e 40).

1

	1	1111111111	22222222	22222222	333333	333333	44444444	44444444	5
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
<i>M. sexta</i>	GTGTT	CGGGAGGTTCC	--GCT	-TTCT	--CG	-T--CACCGAT	-ACTCCTGT		
<i>A. erinacei</i>	GTGTT	TAGGGAGGATCCA	-ATTAT	-CT-	--CGGTGCGT	-CG-	C-GGCGA-GT		
<i>P. communis</i>	GTGTAT	AGGAGGATTCAT	-TTATCCT	-GAGAT	--CA--	TGT-AGCGC-GT			
<i>L. nebulosa</i>	GTGTCAAT	TGAATTC	--TTAT	-CT-	TATATG	-CA--	TAT-GGTAC-GC		
<i>T. paludosa</i>	ATGTAA	AGGA-GTATC	-TA-TTATC	-CTAAAT	-TTAAT	TAT--AT	-GT		
<i>T. regelatio.</i>	GTGTAT	AGGT-AAATC	-T-ATATCT	-TATAT	--GT--	TCC-AATAC-GT			
<i>A. inermis</i>	GTGTAT	-GGA-AAGTC	-TATTAT	-CTGGTGTG	-TA--	TTC-AATAC-GT			
<i>C. cilicrus</i>	GTGTAT	-GGA-AAGTC	-TATTAT	-CTGGTGTG	-TA--	TTC-AATAC-GT			
<i>S. fenestralis</i>	GTGTAT	-GGA-AAGTC	-TATTAT	-CTGGTGTG	-TA--	TTC-AATAC-GT			
<i>D. febrilis</i>	GTGT	-ATGGA-AAGTC	-TA-TTA	-CT--	TTATATAA	AGAT-G	-ATAC-GT		
<i>T. sudeticus</i>	GTGT	-ATAGA-ATATC	-TA-TTATCT	-AGTATAT	TAAAT	-A-ATA--	-GT		
<i>B. cinerea</i>	GTGT	-ATGGG-AAGTC	-TA-TTATC	--TTATAT	GACGGT	-A-GCAC	-GT		
<i>B. coprophila</i>	GTGT	-ATGGA-AAGTC	-TA-TTATC	--TTATG	CAACGAT	-A-ATAT	-GT		
<i>D. melanogas.</i>	GTGT	-ATGGA-GCGTC	-AA-TATCT	-AGTATG	--AG--	-A-A-A--	T		
<i>P. cinerea</i>	GTGT	-ATGGA-GTGC	-CA-TTAACT	CATTACAAT	-AGATG	ATGTAC	-GT		
<i>C. pipiens</i>	GTGTACT	GGA-AATTT	-TG-TCATC	-TGCCATG	--CCCGG	-CGCAGT	-GT		
<i>A. albopictus</i>	GTGTACT	GGA-GAGTT	-CG-TTATC	-TG-CGTA	-ACCGG	TCGCTG	--GT		
<i>S. euryad.</i>	GTGTAC	AGGT-TCCGC	-CA-TTATC	-T--CATAT	TATAAT	TAG-TAT	-GT		
<i>C. tentans</i>	GTGTAA	AGGA-GGGTC	-CA-TTATC	-TTGCACA	-TATAA	CAT-CAT	-AT		

2

	1	1111111111	22222222	22222222	333333	333333	44444444	555555	555555	666666
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234
<i>M. sexta</i>	GTGTT	--C-G-GGAGGTTCCG	-CTTT	--C-T	----	C---G-T-C	-ACCGAT-AC--	TCC-TGT		
<i>A. erinacei</i>	GTGTT	--A-G-GGAGGATCCA	ATTAT	--C-T	----	C-G-G-TGC	-GTCG-C-G---	GCG-AGT		
<i>P. communis</i>	GTG	-T-ATA-GGAGGATTC	ATTATC	-C-T	----	GAG-A-T-C	-AT-G-T-A---	GCG-CGT		
<i>L. nebulosa</i>	GTG	-TC-A-A-TGTGA	ATTCA	-TTAT	-C-T	----	TAT-A-TGC	-AT-A-T-G---	GTA-CGC	
<i>T. paludosa</i>	ATG	-T-AA-A-GGAGTAT	CCTA	-TTAT	-CC-T	--A--	AAT-T-TTT-AA	-T-T-A--	TAT---	-GT
<i>T. regelatio.</i>	GTG	-T-A-TAGGTAAAT	CCTA	-TATT	-C-T	----	TAT-A-TGT	--T-C-C-A---	ATA-CGT	
<i>A. inermis</i>	GTG	-T-A-T-GGAAAGT	CCTA	-TTAT	-C-T	----	GGTGT-TGT	-AT-T-C-A---	ATA-CGT	
<i>C. cilicrus</i>	GTG	-T-A-T-GGAAAGT	CCTA	-TTAT	-C-T	----	GGTGT-TGT	-AT-T-C-A---	ATA-CGT	
<i>S. fenestralis</i>	GTG	-T-A-T-GGAGAGT	CCTA	-CTAT	-C-TC	-A--	TAT-G-TTA	-CG-G-C-A---	ATA-CGT	
<i>D. febrilis</i>	GTG	-T-A-T-GGAAAGT	CCTA	-TTAC	-T-T	----	TAT-A-TAA	-AG-A-T-G---	ATA-CGT	
<i>T. sudeticus</i>	GTG	-T-A-T-AGAATAT	CCTA	-TTAT	-C-T	--A--	GTAT-A-TTA-AA	-T-T-A---	ATA-GTT	
<i>B. cinerea</i>	GTG	-T-A-T-GGGAAGT	CCTA	-TTAT	-C-T	----	TAT-A-TGA	-CG-G-T-A---	GCA-CGT	
<i>B. coprophila</i>	GTG	-T-A-T-GGAAAGT	CCTA	-TTAT	-C-T	----	TAT-G-CAA	-CG-A-T-A---	ATA-TGT	
<i>D. melanogas.</i>	GTG	-T-A-T-GGAGCGT	CAA	-TATT	-C-T	--A--	GTAT-G-AGA-AA	-T-T-A---	ACG--	-AT
<i>P. cinerea</i>	GTG	-T-A-T-GGAGTGT	CCTA	-TTAA	-C-T-CA	--T-T	-A-CAATAG	-A-TGA-T	-GTA	-CGT
<i>C. pipiens</i>	GTG	-T-ACT-GGAAATTTG	-TCAT	-C-T	--GC	-CAT-G--	CC-CG-G--	C---	G-CAGTGT	
<i>A. albopictus</i>	GTG	-T-ACT-GGAGAGT	TCG	-TTAT	-C-T	--G--	CGT-A--	AC-CG-G-T	C---	GCTG-GT
<i>S. euryad.</i>	GTG	-T-ACA-GGTTCGG	CCTA	-TTAT	-C-T	----	CAT-ATTAT	-AA-T-T-A---	GTA-TGT	
<i>C. tentans</i>	GTG	-T-AAA-GGAGGTTCC	ATTAT	-CTT	-G-	-CAC-A-TAT	-AA-A-C-A---	TCA-TAT		

**Figura 15.** Trecho da sequência da região D1 do gene 28S alinhado pelos programas MALIGN (1) e POY (2). Os caracteres autapomórficos estão marcados em azul, enquanto os caracteres que apresentam os cinco estados de caráter, em vermelho.

Foi discutido que o ILD é utilizado como um critério para selecionar parâmetros de alinhamentos baseado na hipótese de minimizar incongruências entre partições de dados (Wheeler & Hayashi, 1998; Terry & Whiting, 2005). Dessa forma, se calcularmos o ILD para os dados dos diferentes programas, podemos ter uma medida da congruência entre as matrizes geradas, podendo escolher, entre os diferentes programas, aquele que apresenta um menor valor de ILD, ou seja, uma maior congruência entre os dados. Vários estudos têm encontrado o mesmo grupo ótimo de parâmetros para as análises filogenéticas, sendo esse o grupo de parâmetros usado nessas análises (1:1:1) (Giribet *et al.*, 2001; Whiting, 2001; Ogden & Whiting, 2003; Svenson & Whiting, 2004; Terry & Whiting, 2005). Calculado o ILD para os resultados da análise combinada das regiões D1 e D7 (Tabela 12), temos que os valores de incongruência são mais altos para os dados analisados com o programa MALIGN, apesar de esse programa gerar os alinhamentos com menor número de caracteres. O POY, mesmo inserindo um grande número de *indels*, apresenta a maior congruência entre os conjuntos de dados, ou seja, os menores valores de ILD.

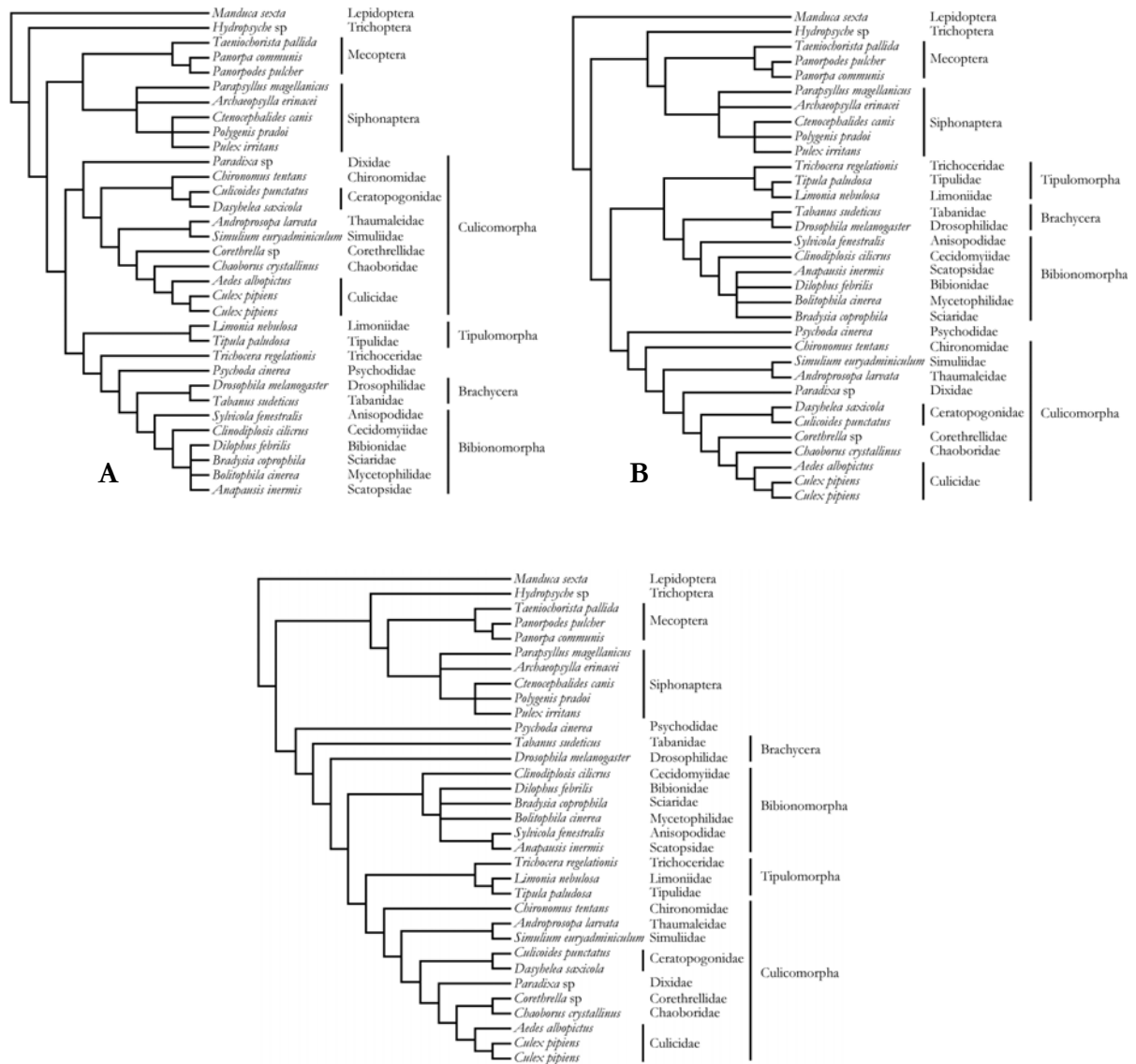
**Tabela 12.** Valores de ILD para os dados analisados com os diferentes programas.

Programa	ILD
POY	0,0232945
CLUSTAL	0,0271816
MALIGN	0,049707

Embora os cladogramas gerados por CLUSTAL e POY tenham comprimentos muito diferentes - os gerados pelo POY são muito mais parcimoniosos - os valores de ILD para esses dois programas são relativamente semelhantes. Terry & Whiting (2005) realizaram um estudo no qual analisaram uma enorme quantidade de parâmetros de custos para esses dois programas, calculando os valores de ILD para cada análise. Segundo esses autores, o conjunto de dados que apresentou valores mais confiáveis foram aqueles analisados pelo programa POY, possivelmente porque os dados são analisados sob um critério de otimização, enquanto o alinhamento de CLUSTAL é um procedimento heurístico simples, baseado em uma única

árvore guia. Além disso, pequenas diferenças de valores de parâmetros também podem ter efeitos drásticos sobre a topologia de uma árvore guia.

Quanto ao efeito dos alinhamentos sobre as hipóteses de relações filogenéticas entre as famílias de Diptera, incluindo todos os táxons terminais (Figura 16), temos que Culicomorpha, grupo considerado muito sólido na literatura, aparece como monofilético. Brachycera aparece como parafilético na análise com o POY e monofilético nos resultados gerados pelos outros dois programas. Bibionomorpha incluindo Scatopsidae aparece como monofilético nas três análises. A família Trichoceridae aparece em Tipulomorpha para as análises de MALIGN e POY e como grupo irmão de Brachycera + Bibionomorpha para a análise de CLUSTAL. Tipulomorpha aparece como grupo irmão de Bibionomorpha + Brachycera (CLUSTAL e MALIGN) ou como grupo irmão de Culicomorpha (POY).





## C

**Figura 16.** Árvores encontradas para as matrizes geradas pelos programas CLUSTAL (A), MALIGN (B) e POY (C), com todos os táxons terminais.

Bibionomorpha resulta como um grupo parafilético em todas as três análises. Como discutido anteriormente, Scatopsidae, que aparece dentro de Bibionomorpha, é atualmente posicionado na infraordem Psychodomorpha (Wood & Borkent, 1989; Courtney, 1990, 1991; Sinclair, 1992; Amorim, 1994; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996). As relações entre os grupos de dípteros basais são bastante variáveis e serão discutidas adiante.

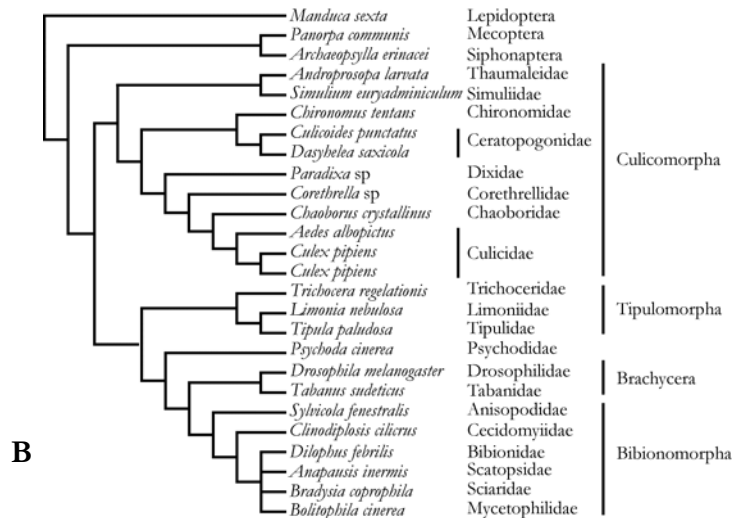
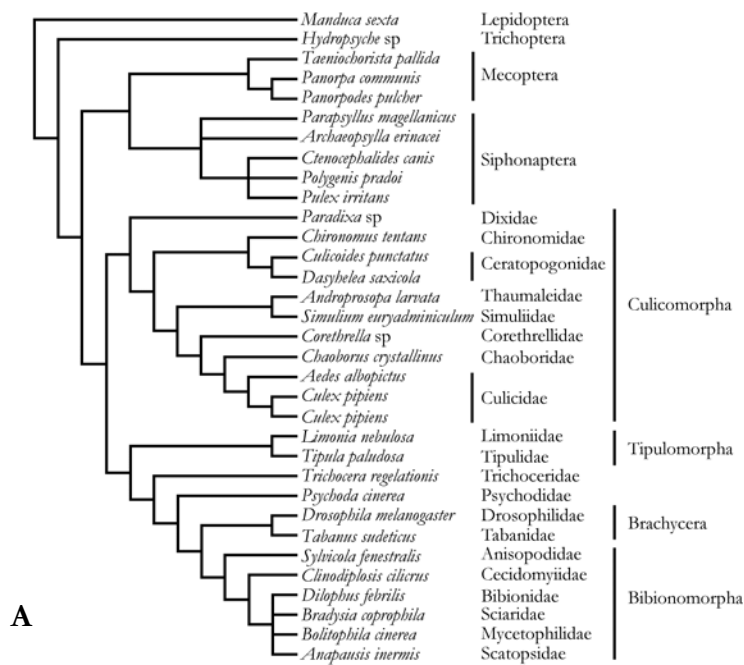
### 4.10. Efeitos da amostragem de táxons sobre as topologias

Rosenberg & Kumar (2001) sugerem que a amostragem de poucos táxons terminais em análises filogenéticas de grandes grupos não penaliza a acurácia da reconstrução de árvores e que a redução da amostragem não é um problema para análises filogenéticas. Entretanto, dois outros trabalhos colidem diretamente com os resultados obtidos por esses autores e concluem que o aumento na amostragem de táxons terminais tem um grande efeito positivo no aumento da acurácia das análises (Pollock *et al.*, 2002; Zwickl & Hillis, 2002). Dessa forma, foram procedidas análises retirando táxons terminais dos grupos interno e externo, verificando o efeito dessa diminuição na amostragem sobre as topologias geradas.

#### 4.10.1. Retirada de táxons terminais do grupo externo

Com a retirada de sete dos dez táxons terminais do grupo externo nas análises, a topologia resultante da análise dos dados com o programa CLUSTAL sofreu pouca alteração (Figura 17). Com um menor número de táxons enraizando a topologia, as relações internas de Culicomorpha sofreram algumas mudanças, sendo que o clado Thaumaleidae + Simuliidae

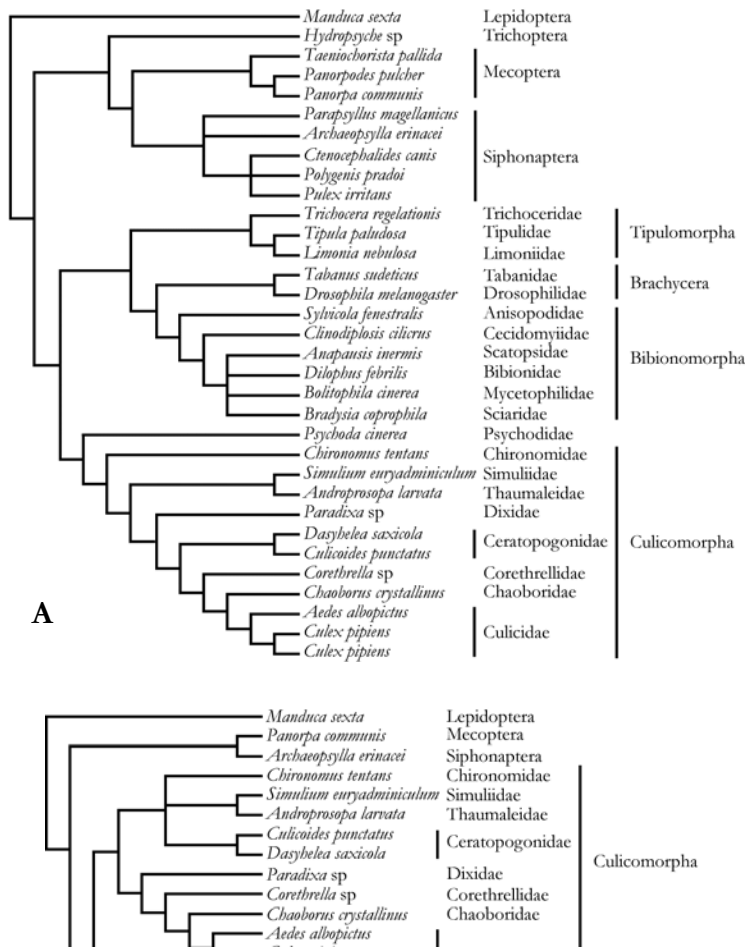
ficou como o mais basal do grupo ao invés de Dixidae. Nenhum desses dois resultados é concordante com as filogenias propostas para Culicomorpha (Saether, 2000). Trichoceridae aparece dentro de Tipulomorpha, mas as relações entre os grandes grupos continuam as mesmas, com Culicomorpha como grupo irmão do restante de Diptera.



**Figura 17.** Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa com a retirada de sete dos dez táxons terminais do grupo externo (B).

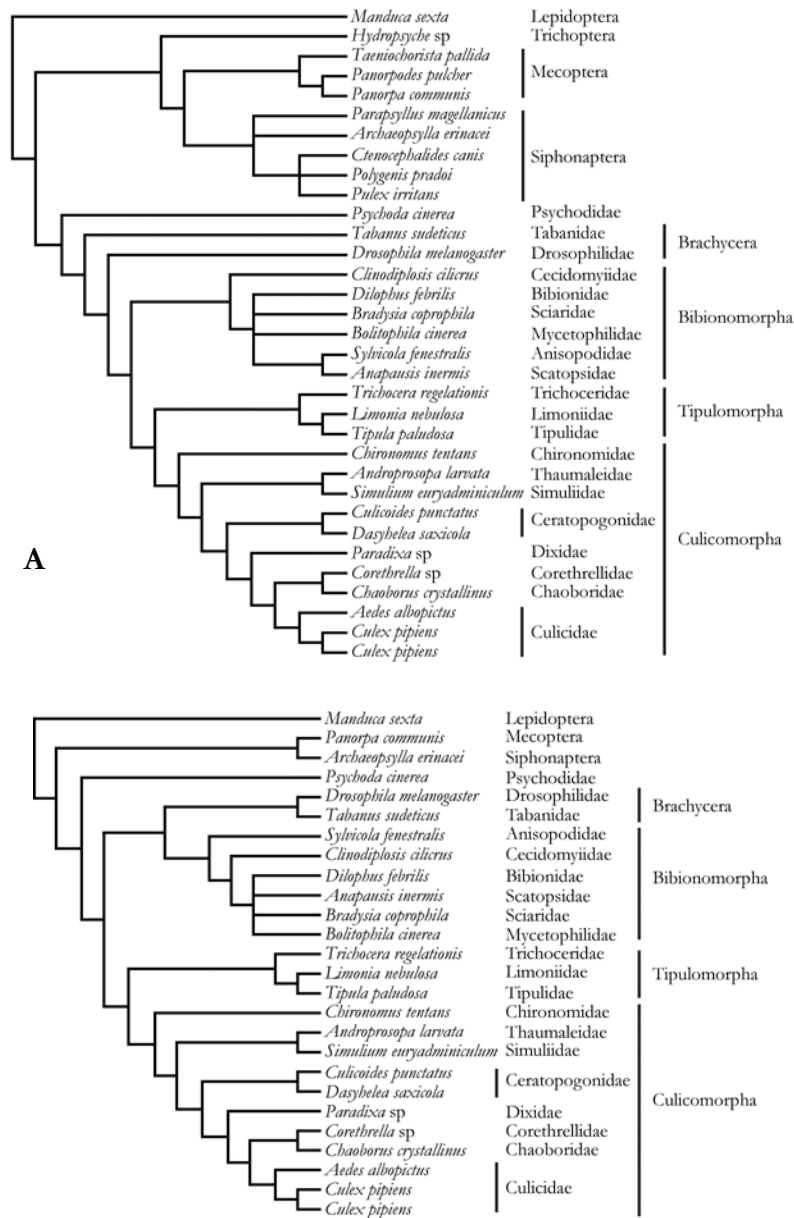
Os resultados das análises de MALIGN sofreram mais alterações com a retirada dos táxons externos (Figura 18). Psychodidae, que antes aparecia como grupo irmão de Culicomorpha, aparece como grupo irmão de Tipulomorpha + Trichoceridae. A mudança de Psychodidae transfere Culicomorpha para a base de Diptera. Brachycera aparece como parafilético, mas permanece próximo de Bibionomorpha, com Tabanidae na base do clado formado por Drosophilidae + Bibionomorpha (incluindo Scatopsidae). As relações internas de Bibionomorpha e Culicomorpha também sofrem alterações, demonstrando a sensibilidade desses cladogramas para esse algoritmo de alinhamento.

Os resultados gerados pelo POY (Figura 19) sofreram pouca interferência com a retirada dos táxons do grupo externo. Nas topologias geradas com esse programa, Psychodidae continua aparecendo como grupo mais basal de Diptera. Trichoceridae continua incluído na base de Tipulomorpha e as relações internas em Culicomorpha ficam inalteradas, bem como a relação filogenética entre esses dois grupos. Brachycera, que aparecia como parafilético, resulta como um grupo monofilético, irmão de Bibionomorpha + Scatopsidae. Ocorrem mudanças nas relações internas de Bibionomorpha.



**B**

**Figura 18.** Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa com a retirada de sete dos dez táxons terminais do grupo externo (B).



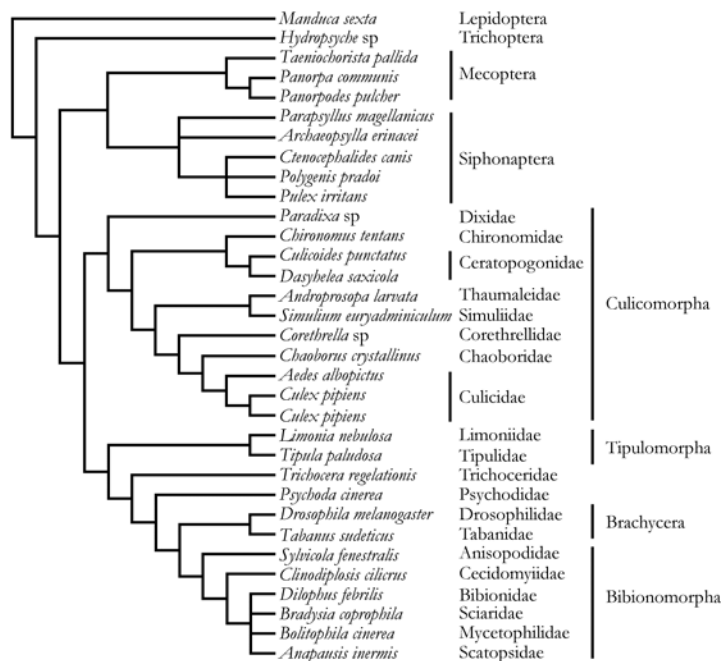
## B

**Figura 19.** Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo externo (B).

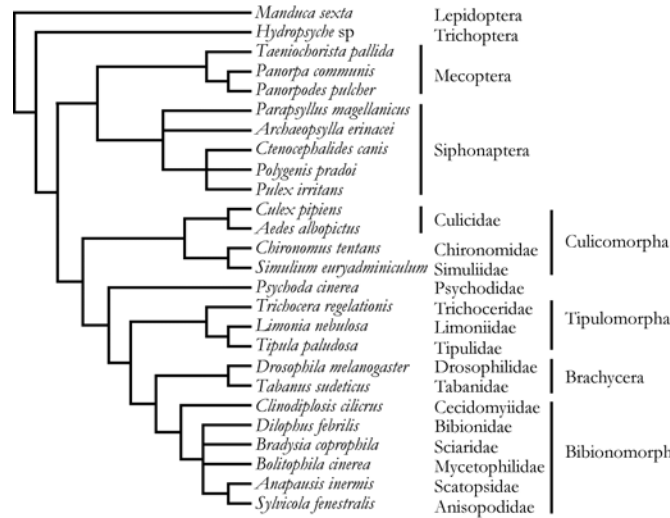
### 4.10.2. Retirada de táxons terminais do grupo interno

Todos os terminais retirados nessa parte da análise pertencem aos Culicomorpha. A diminuição do número de terminais desse grupo não afetou em nenhum caso a monofilia sob esses parâmetros de custos (1:1:1). No entanto, as relações entre as três famílias que permaneceram varia na análise. Somente a topologia encontrada pelo matriz gerada por CLUSTAL (Figura 20) recuperou a hipótese de monofilia de Chironomoidea. Nas outras duas topologias (Figuras 21 e 22) Simuliidae aparece como grupo irmão de Culicidae, tendo Chironomidae na base desse clado. Chironomoidea é considerado parafilético por Saether (2000), mas pela exclusão a família Thaumaleidae e não pela união desses dois grupos.

A retirada de táxons terminais de Culicomorpha, no entanto, produziu algumas alterações relevantes nas relações entre os demais Diptera nas análises realizadas pelo CLUSTAL (Figuras 20 A-B) não sofreram muitas variações. Trichoceridae, que antes aparecia na base de um clado formado por Psychodidae, Bibionidae (incluindo Scatopsidae) e Brachycera, aparece em Tipulomorpha.



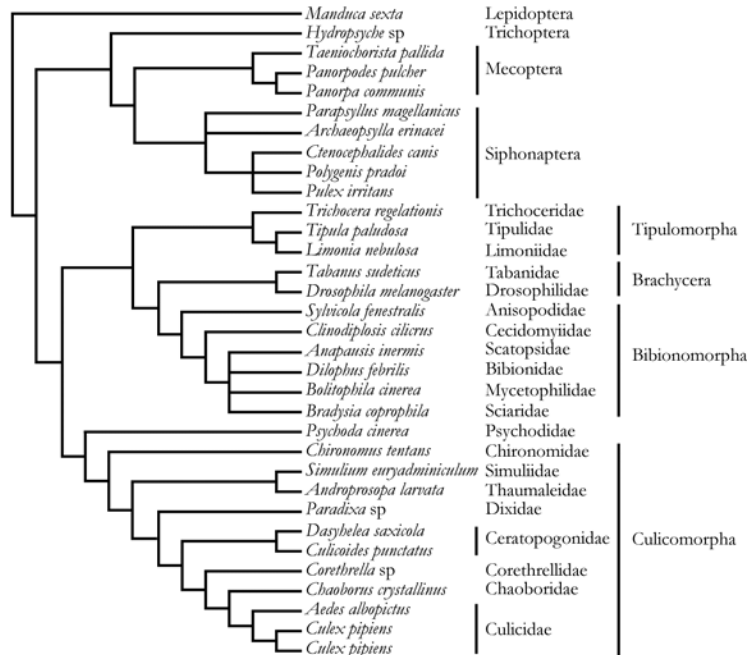
A



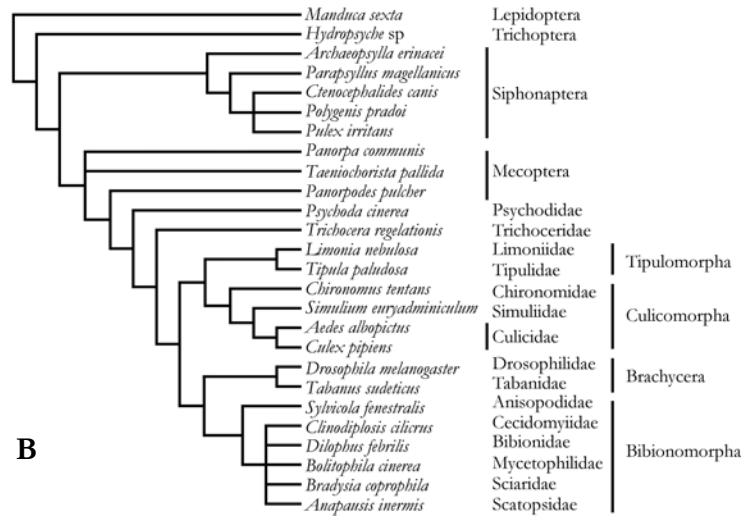
B

**Figura 20.** Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).

Com o MALIGN, a retirada de elementos do grupo interno gera um efeito interessante, Mecoptera tornando-se parafilético em relação à Diptera (Figura 21). Psychodidae aparece como grupo irmão de todos os outros Diptera. Bibionomorpha se mantém como grupo irmão de Brachycera e Tipulomorpha aparece como grupo irmão de Culicomorpha. Essa topologia é muito semelhante à gerada por POY (Figura 22), com a diferença que Trichoceridae aparece dentro de Tipulomorpha, e para MALIGN como grupo irmão de todos os outros Diptera menos Psychodidae.



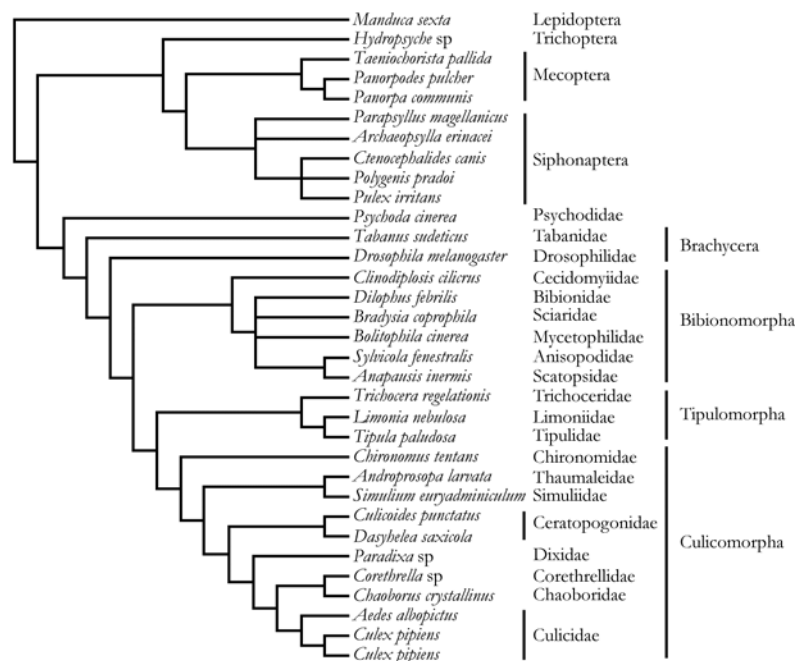
A

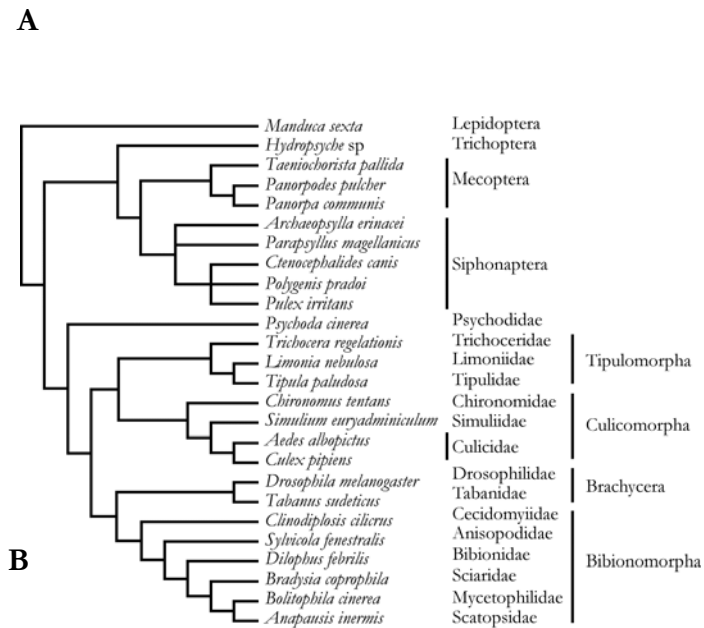


B

**Figura 21.** Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).

As relações entre os grupos de Diptera recuperadas pelo programa POY (Figura 22B) com a diminuição de terminais de Culicomorpha é exatamente a mesma que para a análise com redução de terminais do grupo externo (Figura 19B). O clado formado por Brachycera e Bibionomorpha + Scatopsidae aparece como grupo irmão de Culicomorpha + Tipulomorpha (incluindo Trichoceridae). Psychodidae aparece como grupo irmão do restante de Diptera.





**Figura 22.** Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa com a retirada de táxons do grupo interno (B).

#### 4.11. Efeito da amostragem de diferentes regiões do gene sobre as topologias

A análise combinada de diferentes dados pode minimizar erros de amostragem taxonômica e maximizar o poder explanatório de dados filogenéticos, já que a independência dos caracteres aumenta a quantidade de evidências e pode ser usada para escolher entre hipóteses evolutivas alternativas (Kluge, 1989; Kluge & Wolf, 1993).

##### 4.11.1. Análise da região D7 do gene 28S

As análises da região D7, também incluem a retirada de táxons do grupo interno e do grupo externo, já que em todas as análises anteriores foram utilizadas seqüências dessa região. Aqui há uma diminuição no número de terminais, restando somente aqueles utilizados por Friedrich & Tautz (1997a).

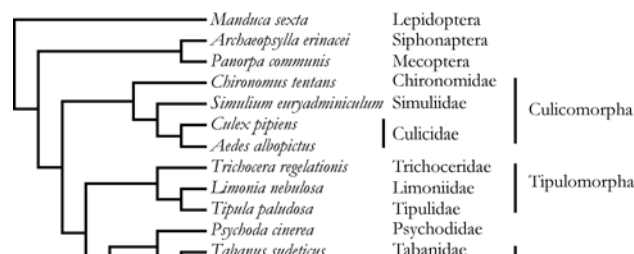
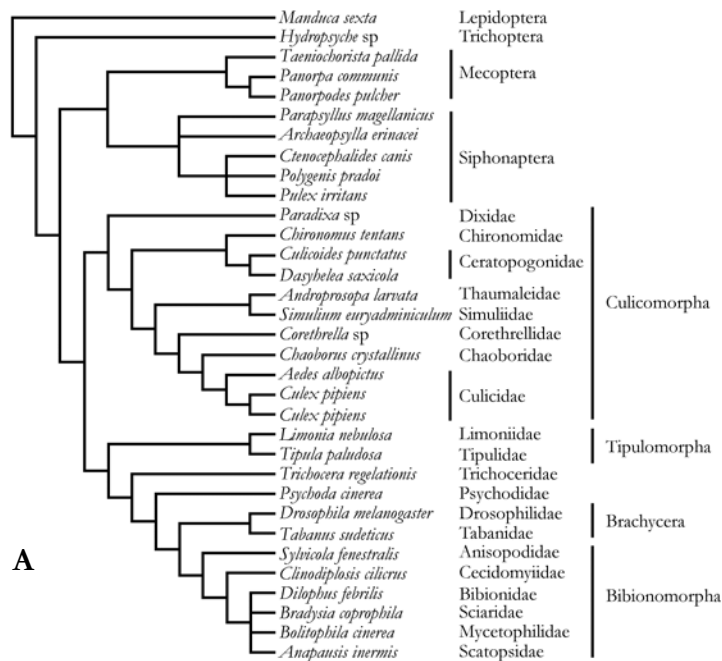
Para a topologia encontrada para o alinhamento resultante de CLUSTAL (Figura 23), ocorrem algumas mudanças relevantes entre as relações dos grupos de Diptera. Culicomorpha continua como o clado mais basal. As diferenças encontradas são que Trichoceridae é incluído



em Tipulomorpha e Psychodidae forma um clado com Brachycera, que antes resultava como grupo irmão de Bibionomorpha mais Scatopsidae.

O programa MALIGN gerou três alinhamentos igualmente parcimoniosos para esse conjunto de terminais (Figura 24). Dois desses alinhamentos resultam em topologias idênticas (Figura 24C - D), que diferem em vários pontos da topologia gerada pelo outro alinhamento (Figura 24B) e da topologia inicial (Figura 24; A). A posição de Psychodomorpha é variável, aparecendo ora como grupo mais basal de Diptera ora como grupo irmão de Culicomorpha e ora como grupo irmão de Brachycera + Bibionomorpha (incluindo Scatopsidae). Trichoceridae resulta dentro ou fora de Tipulomorpha, mas nunca formando um clado com Psychodidae.

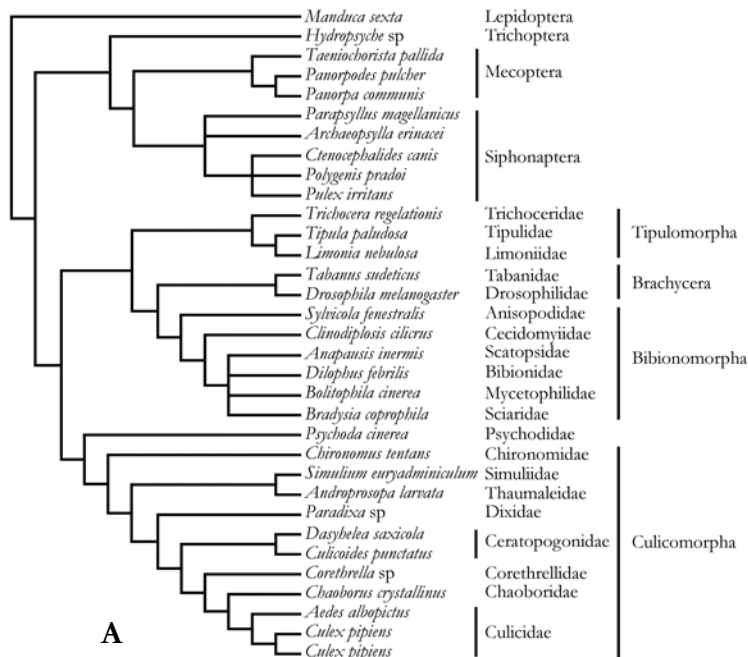
Os cladogramas gerados pelos três alinhamentos são igualmente parcimoniosos e uma análise de consenso geraria um cladograma quase que totalmente não resolvido para as relações dos grandes grupos de Diptera. O único ponto de concordância entre esses resultados é a monofilia do clado Bibionomorpha (incluindo Scatopsidae) + Brachycera.

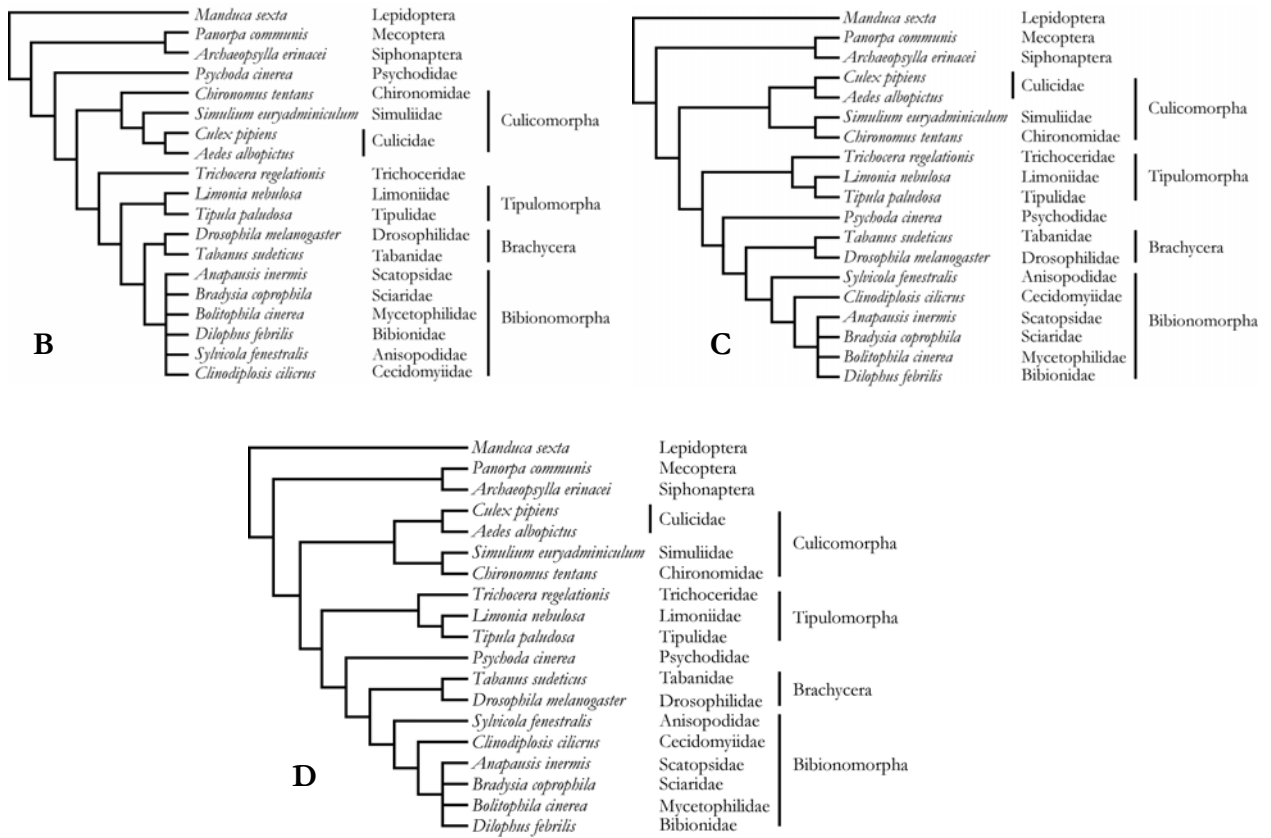


## B

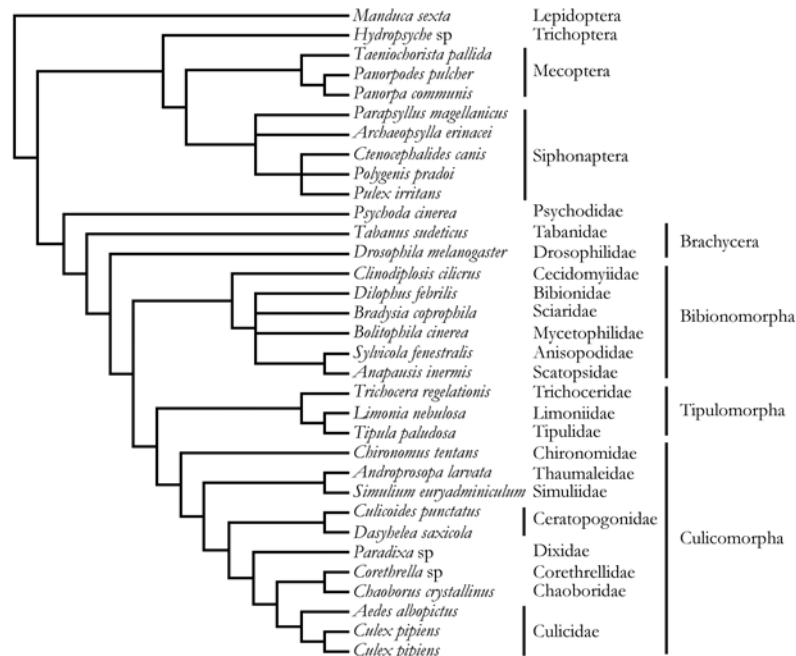
**Figura 23.** Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa para a região D7 do gene 28S (B).

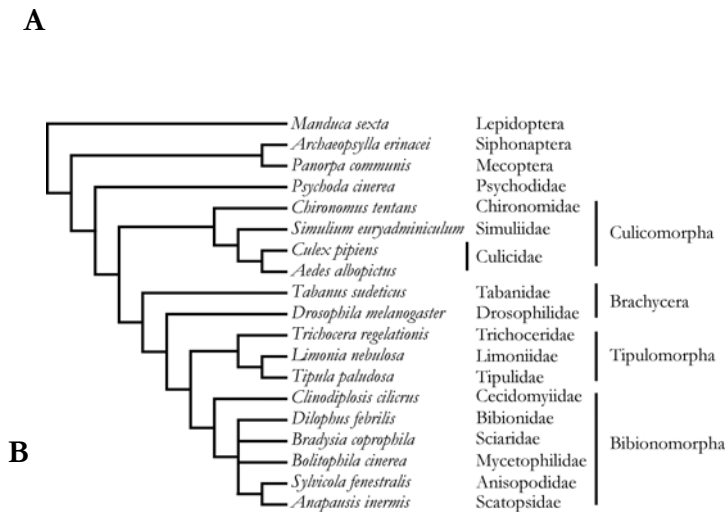
As diferenças de topologia geradas pelo POY para a região D7 do gene 28S são muito grandes (Figura 25). O único ponto de concordância com a topologia original é a posição basal de Psychodidae com relação ao restante de Diptera. Culicomorpha, que antes resultava como grupo irmão de Tipulomorpha (incluindo Trichoceridae), aparece agora na base do clado formado por esse infraordem, juntamente com Bibionomorpha (mais Scatopsidae) e Brachycera, que resulta parafilético. Esse posicionamento basal de Psychodidae nunca foi sugerido por qualquer estudo de dados morfológicos.





**Figura 24.** Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e as encontradas pelo mesmo programa para a região D7 do gene 28S (B, C e D).





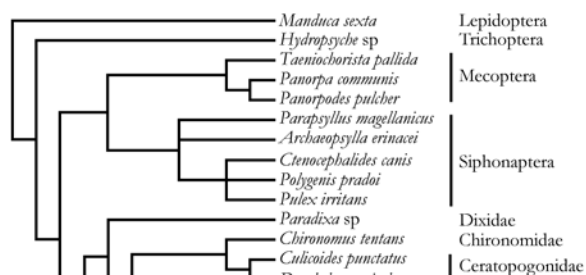
**Figura 25.** Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada pelo mesmo programa para a região D7 do gene 28S (B).

#### 4.11.2. Análise da região D1 do gene 28S

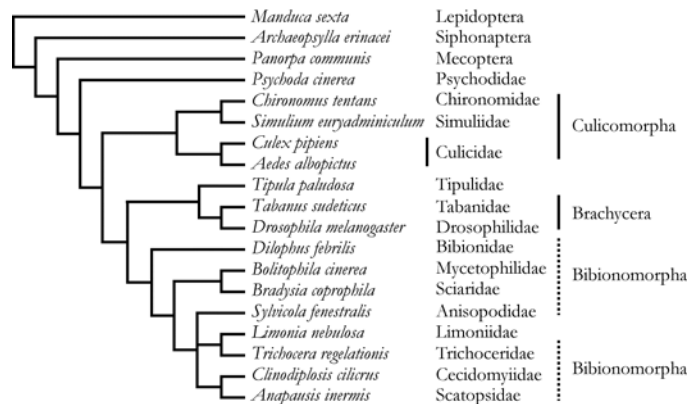
Quando analisamos somente os dados da região D1 do gene 28S, as topologias sofrem muita variação quando comparadas às análises iniciais. Essas diferenças indicam a incongruência com a região D7 do mesmo gene. Nas topologias geradas pelos três programas, somente Brachycera e Culicomorpha aparecem como grupos monofiléticos, sendo que as outras infraordens (Tipulomorpha, Bibionomorpha e Psychodomorpha) resultam como grupos parafiléticos.

Na análise com CLUSTAL (Figura 26), Brachycera aparece como grupo irmão de Tipulidae, e Tipulomorpha resulta parafilético. Limoniidae, formando juntamente com Trichoceridae um grupo monofilético, é incluído em Bibionomorpha, que antes já aparecia como um grupo parafilético pela inclusão de Scatopsidae. Pela primeira vez para os resultados gerados por esse programa, Psychodidae é posicionado na base de Diptera. Devido à perda da monofilia dos grandes grupos, a discussão dos dados se torna difícil, já que os resultados apresentados são profundamente discordantes de qualquer topologia apresentada anteriormente.

Fica clara a incongruência entre esses dois conjuntos de dados (regiões D1 e D7), mesmo sendo trechos de um mesmo gene.



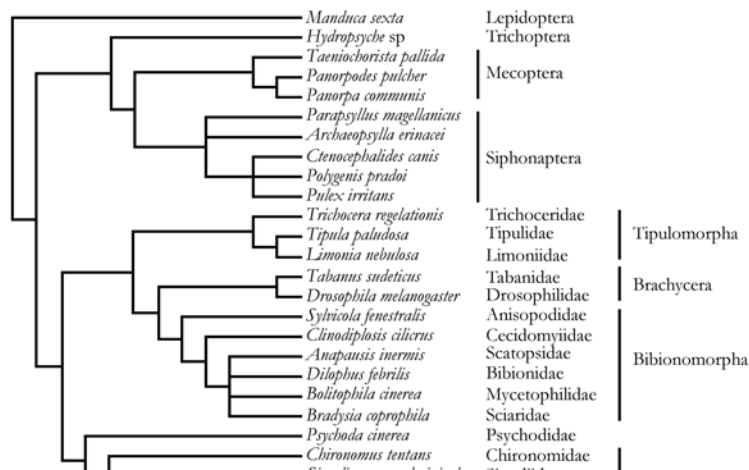
A

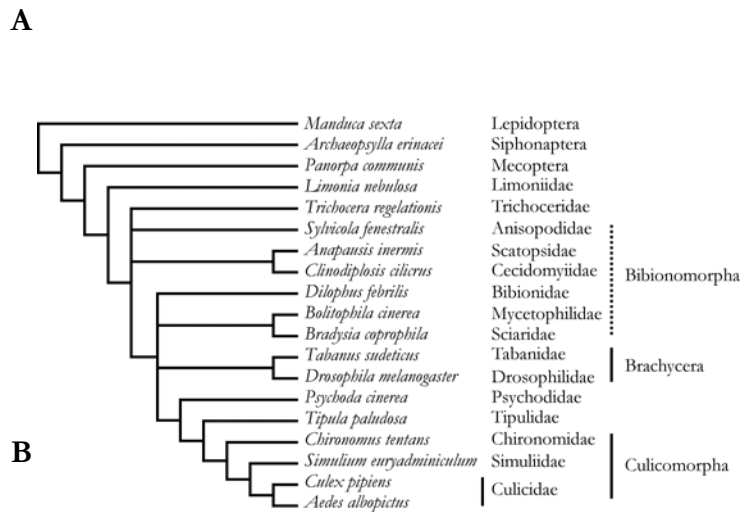


B

**Figura 26.** Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada pelo mesmo programa para a região D1 do gene 28S (B).

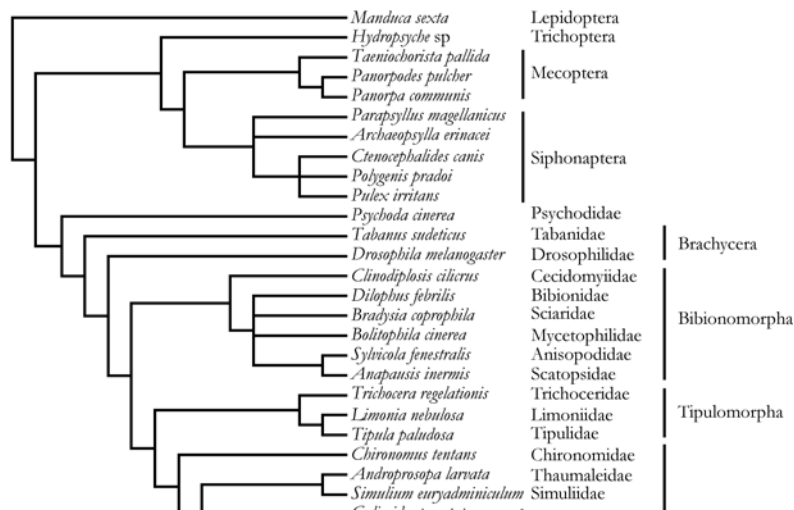
A topologia gerada por MALIGN apresenta várias politomias (Figura 27). Limoniidae aparece isolado como grupo mais basal de Diptera, resultado ainda não encontrado nas análises realizadas nesse trabalho. Bibionomorpha resulta em uma politomia, juntamente com Trichoceridae e Brachycera. Psychodidae aparece como grupo irmão de Tipulidae + Culicomorpha.

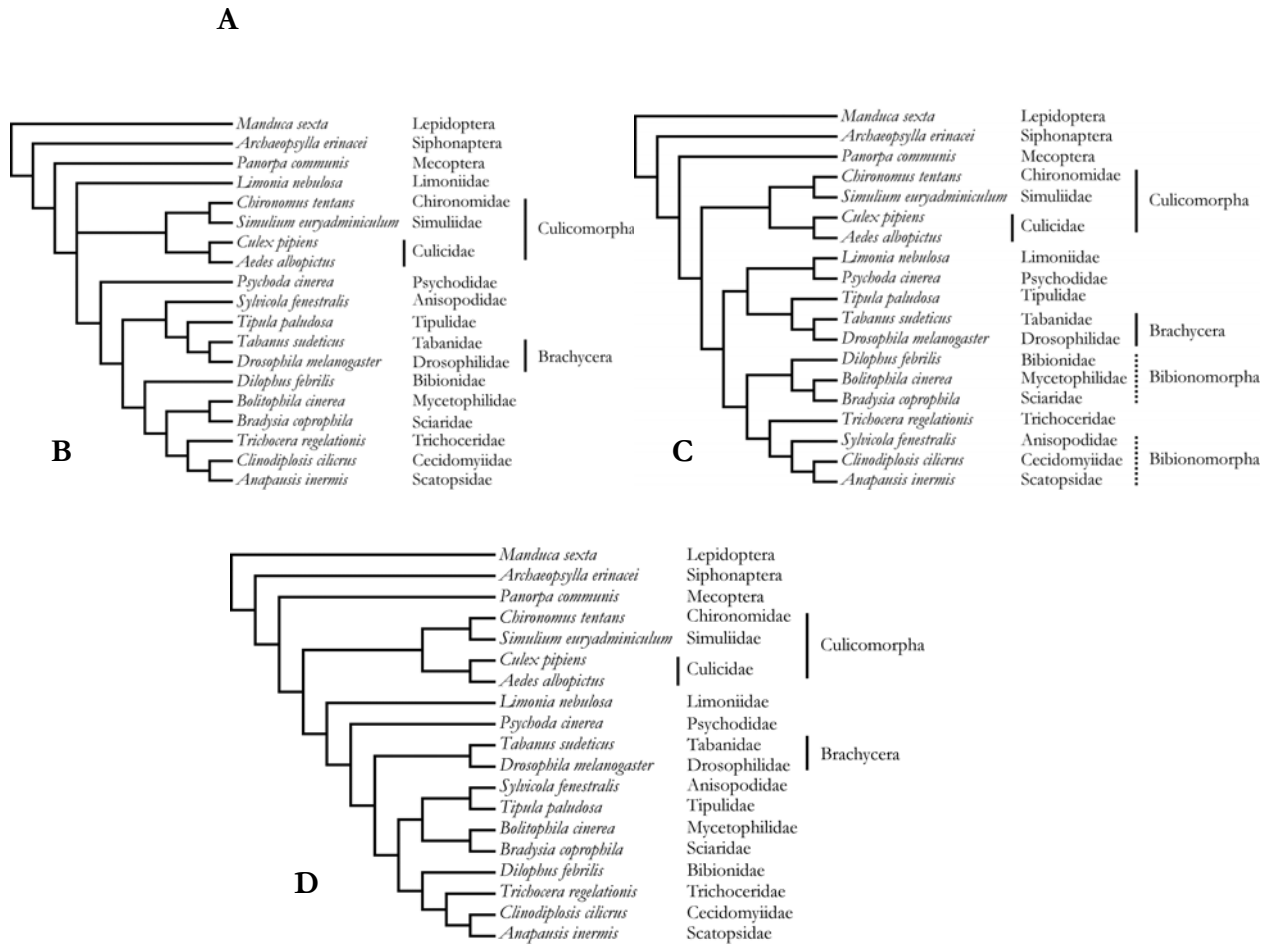




**Figura 27.** Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e a encontrada pelo mesmo programa para a região D1 do gene 28S (B).

As árvores encontradas pelos três alinhamentos igualmente parcimoniosos gerados pelo POY diferem muito entre si (Figura 28B - D) e, exceto pela monofilia de Culicomorpha, diferem também da árvore inicial. Os poucos pontos de concordância entre as topologias são as relações filogenéticas de Culicomorpha, bem como sua monofilia, a monofilia de Brachycera, a união das famílias Mycetophilidae e Sciaridae num clado monofilético e a junção de Scatopsidae + Cecidomyiidae.



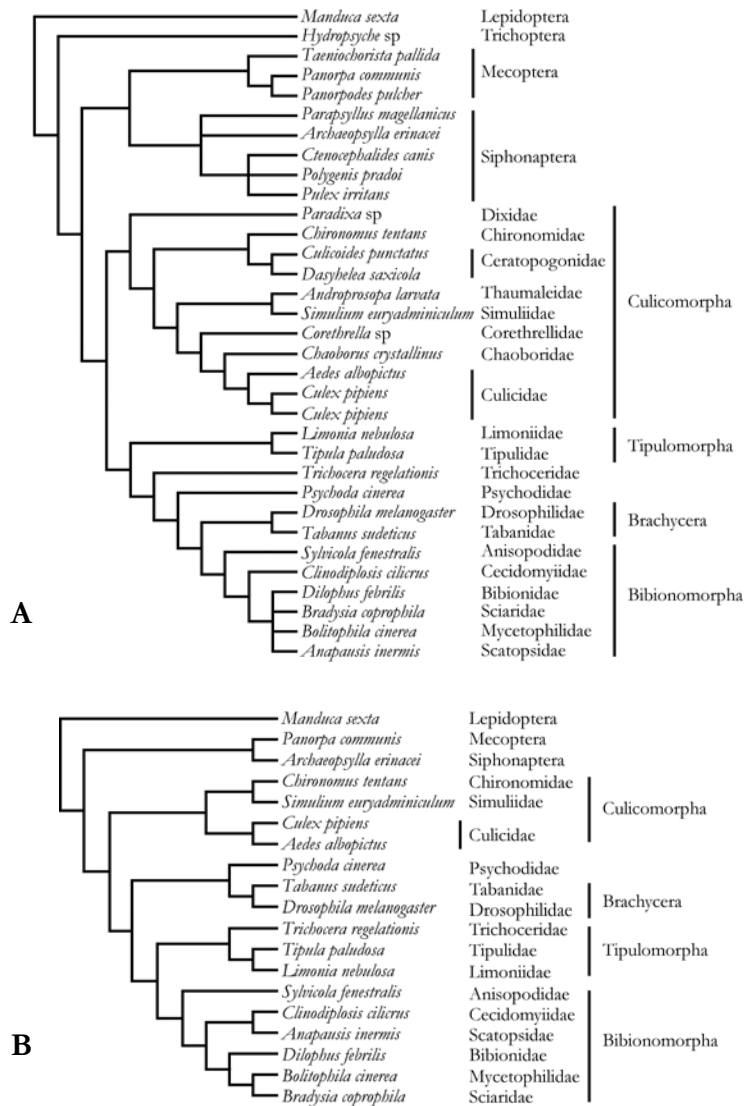


**Figura 28.** Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e as encontradas pelo mesmo programa para a região D1 do gene 28S (B, C e D).

### 4.11.3. Análise combinada das regiões D1 e D7 do gene 28S

Ao analisarmos conjuntamente os dois grupos de dados (regiões D1 e D7), novamente as monofilias dos grandes grupos de Diptera são recuperadas, na maior parte dos casos.

A topologia encontrada para os alinhamentos feitos por CLUSTAL para os dados combinados analisados mostra relações filogenéticas dos grandes grupos de Diptera diferente das demais encontradas com esse programa (Figura 29).



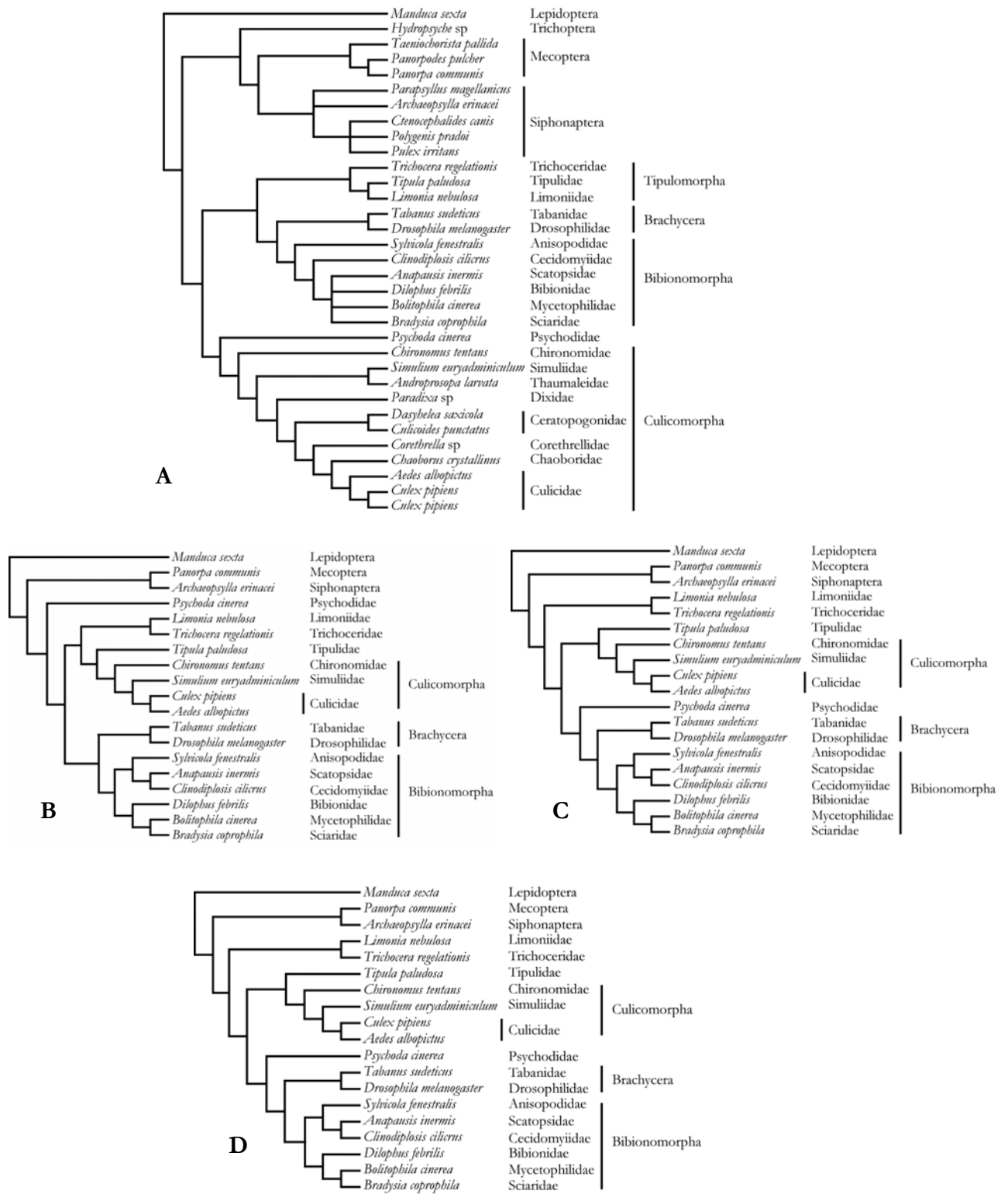
**Figura 29.** Árvore encontrada para a matriz inicial gerada pelo programa CLUSTAL (A) e a encontrada em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B).



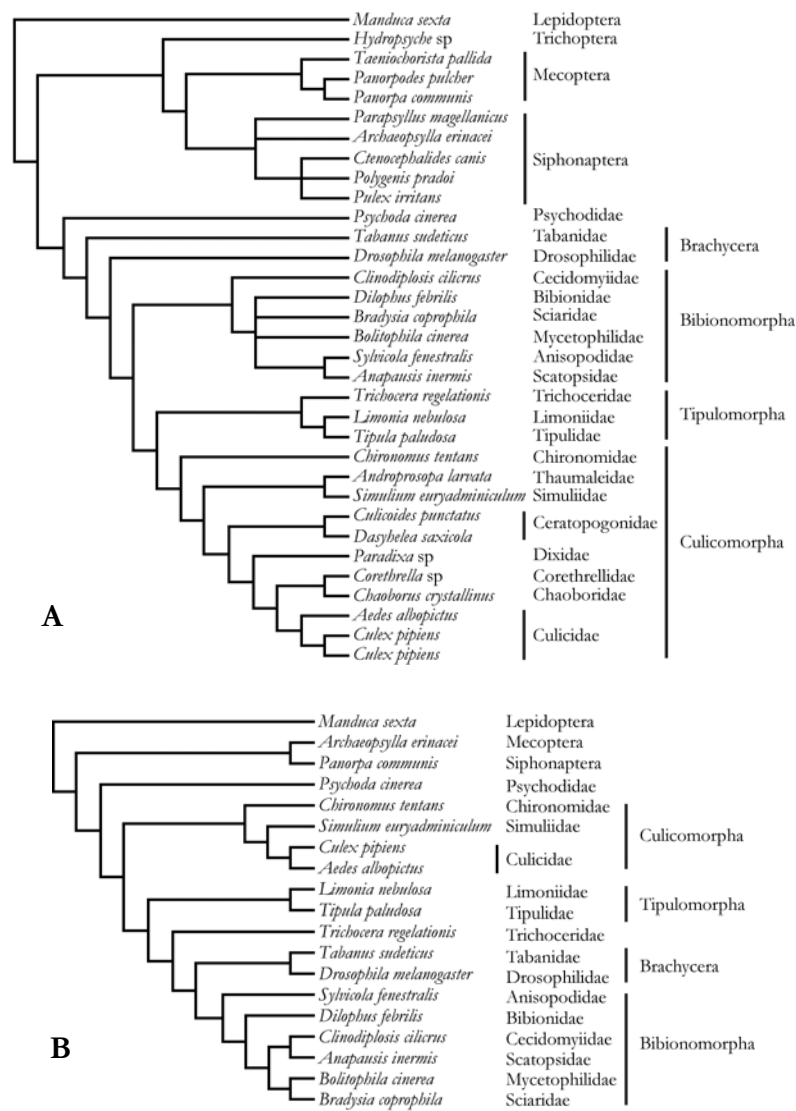
Culicomorpha, como para todos os outros resultados (exceto para a região D1), aparece como grupo mais basal de Diptera. O clado monofilético Psychodidae + Brachycera é recuperado (Figura 29B), como no caso da filogenia gerada para a região D7. A união de Psychodomorpha + Brachycera é sugerida por vários autores (Sinclair, 1992; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996), mas aqui não podemos dizer que exatamente esse resultado foi obtido, pois duas famílias incluídas em Psychodomorpha por esses mesmos pesquisadores, Trichoceridae e Scatopsidae, aparecem em Tipulomorpha e Bibionomorpha, respectivamente, recuperando a filogenia proposta por Hennig (1973), pouco aceita atualmente.

Novamente dois dos três cladogramas gerados pelos alinhamentos resultantes do programa MALIGN foram idênticos (Figura 30C-D). Os cladogramas são concordantes quanto à monofilia dos grupos Brachycera + Bibionomorpha (incluindo Scatopsidae), Limoniidae + Trichoceridae e Tipulidae + Culicidae. Ao contrário dos resultados de CLUSTAL, a monofilia de Tipulomorpha não foi recuperada por esse conjunto de dados. A congruência entre dados gerados por MALIGN, inferida por ILD, é a menor dentre os três programas de alinhamento o que pode ser claramente observado nas topologias, já que, para os outros dois programas, os grupos monofiléticos que aparecem nas primeiras análises são recuperados.

Os alinhamentos gerados pelo POY recuperam a monofilia dos grupos da primeira análise e ainda recuperam a monofilia de Brachycera, que não aparece na análise com todos os terminais (Figura 31). As relações filogenéticas entre os grandes grupos de Diptera também não havia aparecido em nenhum dos resultados anteriores com o POY. Como para todos os outros resultados, Psychodidae é posicionado como grupo mais basal de Diptera. A união de Brachycera + Bibionomorpha (incluindo Scatopsidae) só não é proposta na análise da região D7, na qual Tipulomorpha aparece como grupo irmão desse último clado.



**Figura 30.** Árvore encontrada para a matriz inicial gerada pelo programa MALIGN (A) e as encontradas em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B, C e D).



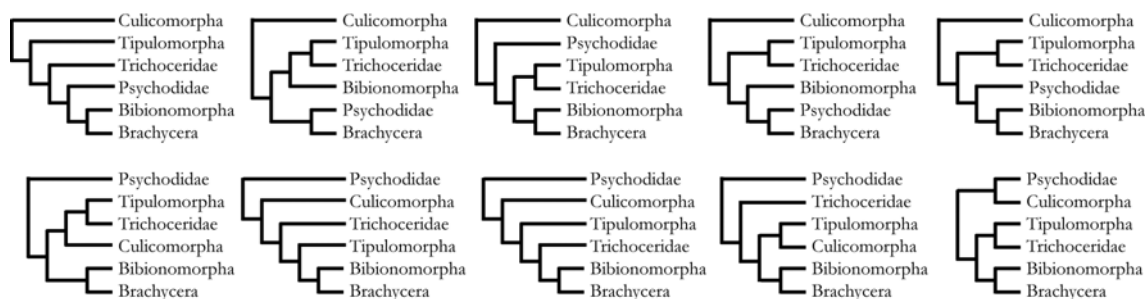
**Figura 31.** Árvore encontrada para a matriz inicial gerada pelo programa POY (A) e a encontrada em análise combinada das regiões D1 e D7 do gene 28S, pelo mesmo programa (B).

#### 4.12. A influência dos diferentes programas de alinhamento

Em todas as análises realizadas para o gene nuclear 28S com os diferentes programas de alinhamento, obtivemos um total de 24 topologias que, exceto em poucos casos, foram diferentes umas das outras. Exceto os casos em que os alinhamentos de MALIGN produziram topologias idênticas, encontramos outros dois casos de concordância entre as relações

filogenéticas dos grandes grupos de Diptera. As análises filogenéticas dos alinhamentos gerados por CLUSTAL e MALIGN, sob diferentes condições de análise (retirada de táxons do grupo externo para CLUSTAL e análise da região D7 para MALIGN), resultaram em topologias que recuperam as mesmas relações entre as infraordens. O mesmo acontece com dois alinhamentos gerados por POY (retirada de táxons do grupo externo e retirada de táxons do grupo interno).

Analisando todas as topologias resultantes, verificamos que elas possuem poucos pontos de concordância. Além disso, as relações entre os grupos de Diptera concordam em muitos poucos casos com as hipóteses filogenéticas já existentes. Considerando somente os cladogramas em que temos como grupos monofiléticos Tipulomorpha, Brachycera, Culicomorpha, Bibionomorpha incluindo Scatopsidae, e as famílias Psychodidae e Trichoceridae, temos um conjunto de 10 topologias que demonstram as diferentes relações filogenéticas entre esses grupos (Figura 32).



**Figura 32.** Conjunto de topologias que representam as diferentes hipóteses de relações filogenéticas entre os grupos de dípteros basais encontradas no presente estudo.

Com os resultados resumidos aqui, é fácil perceber que, para todo o conjunto de dados, temos dois táxons que mais frequentemente aparecem na base de Diptera: Culicomorpha e Psychodomorpha. Oosterbroek & Courtney (1995), estudando os grupos basais de Diptera, sugerem que um clado formado por Culicomorpha + Ptychopteromorpha seja grupo irmão do restante de Diptera. A análise de parcimônia de transversões de Friedrich & Tautz (1997a) também recupera esse posicionamento para Culicomorpha.

A hipótese de Psychodidae como grupo basal de Diptera é inédita e pouco provável. Dados morfológicos reúnem esse grupo com Blephariceromorpha, Ptychopteromorpha + Culicomorpha (Courtney, 1990, 1991; Wood & Borkent, 1989) ou Brachycera (Sinclair, 1992; Oosterbroek & Courtney, 1995; Amorim *et al.*, 1996) e não aparecem posicionados nem

mesmo como um dos mais basais. A infraordem Psychodomorpha (*sensu* Wood & Borkent, 1989) resulta como parafilética em todas as análises. Trichoceridae e Scatopsidae não são posicionados junto com Psychodomorpha em qualquer uma das topologias. Trichoceridae aparece frequentemente com Tipulomorpha e Scatopsidae sempre com Bibionomorpha. O posicionamento dessas famílias dentro desses grupos é concordante com Hennig (1981), mas a inclusão de Scatopsidae em Bibionomorpha contradiz caracteres morfológicos abundantes de larvas e de adultos.

A infraordem Tipulomorpha *sensu* Hennig (1973, 1981) é freqüentemente recuperada nas várias topologias encontradas, entretanto seu posicionamento na base de Diptera não aparece em nenhum dos resultados. Em contraste, esse grupo é encontrado em alguns casos como grupo irmão de Bibionomorpha + Brachycera, uma conclusão bastante ímpar em relação às demais fontes de evidência disponíveis.

Bibionomorpha, na maioria dos resultados aparece como grupo irmão de Brachycera – um táxon chamado, na literatura, de Neodiptera. Podemos dizer que esse é um dos poucos pontos de concordância da análise de dados moleculares realizada aqui com as hipóteses previamente estabelecidas por dados morfológicos. Apesar de haver discordância sobre o posicionamento desse grupo em Diptera, alguns autores realmente o posicionam como o grupo mais apical de dípteros nematóceros em relação aos braquíceros (Hennig, 1981; Amorim, 1993; Michelsen, 1996).

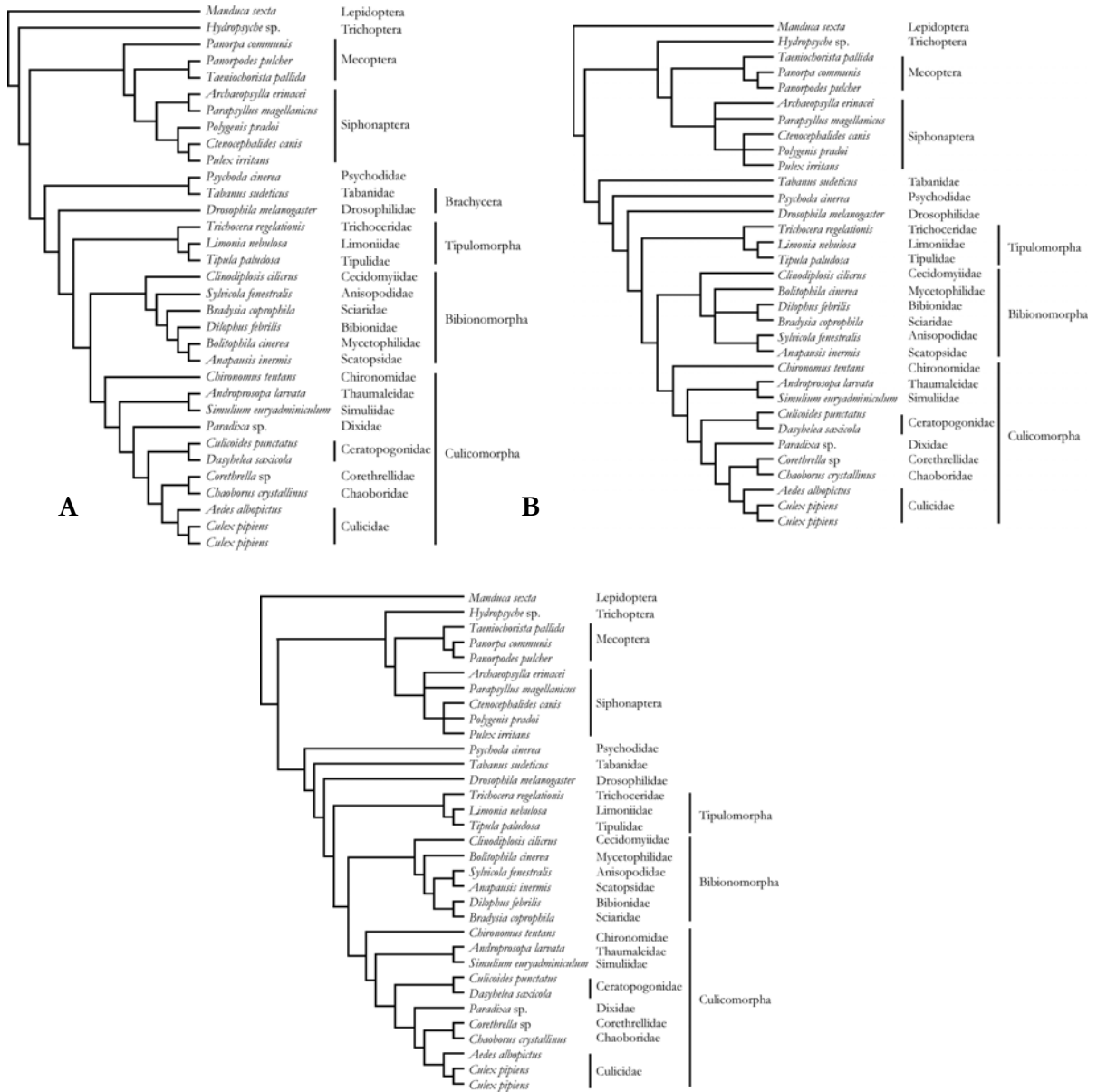
As diversas discordâncias entre os resultados encontrados, mostram que, mesmo com um enraizamento mais confiável ou um número um pouco maior de táxons terminais, não foi possível estabelecer uma maior congruência entre as filogenias, ou seja, não há robustez nas análises e, portanto, qualquer das topologias deve ser vista com suspeição. Isso mostra, portanto, que a incongruência entre as topologias obtidas deve ter sido causada por amostragem deficiente de espécies dentro dos gêneros, de gêneros dentro das famílias e de famílias dentro de cada ordem.

#### **4.13. Diferentes algoritmos de análise**

É muito comum encontrarmos trabalhos sobre relações filogenéticas de diferentes grupos com base em dados moleculares nos quais diferentes algoritmos de análise são utilizados e seus resultados comparados (e.g. Gimeno *et al.*, 1997; Kambhampati & Charlton,

1999; Sallum *et al.*, 2002). As análises de Maximum likelihood e Máxima Parcimônia utilizadas por Friedrich & Tautz (1997a) apresentam muitos pontos de discordância quanto às relações filogenéticas encontradas para os grandes grupos de Diptera. De fato, as variações entre as topologias geradas dependendo do algoritmo de análise utilizado, são resultados das diferenças entre as premissas de análise dos vários algoritmos.

Os cladogramas pelos métodos de Máxima Parcimônia e Maximum likelihood foram idênticos, enquanto que o cladograma obtido pela análise de Neighbor-joining difere dos outros dois somente pela união de Psychodidae e Tabanidae em um clado monofilético (Figura 33). Contudo, em todos eles, Brachycera aparece na base ou próximo à base de Diptera, possivelmente indicando uma inversão do sinal filogenético ao menos nesse nível de análise.



## C

**Figura 33.** Cladogramas gerados pelos diferentes métodos de análise: (A) Neighbor-joining; (B) Maximum likelihood; e (C) Máxima Parcimônia, para a região D7 do gene 28S.

### 5. CONSIDERAÇÕES FINAIS

- Mesmo com o virtual esgotamento de táxons e genes comparáveis para Diptera existentes no GenBank, as relações filogenéticas para os grupos basais dessa ordem permanecem incertas;
- Para o nível taxonômico estudado aqui, os genes nucleares (18S e 28S) foram os melhores indicadores para as relações filogenéticas de Diptera no quanto a comparação com diferentes reconstruções utilizando dados morfológicos é confiável;
- A subamostragem de táxons dos grupo externo e interno, aparentemente, pode ser responsável pelas diferenças de resultados entre resultados obtidos com dados morfológicos e moleculares nas análises com os genes 18S e 28S;
- Os genes nucleares utilizados neste estudo apresentam maior resolução para as relações filogenéticas de Diptera que os genes mitocondriais;
- Apesar de os genes mitocondriais (12S, 16S, COI e COII) proporcionarem uma melhor amostragem de famílias de Diptera, eles não fornecem sinal filogenético apropriado para os níveis taxonômicos abordados neste estudo;
- Os diferentes programas de alinhamento utilizados para seqüências do gene nuclear 28S geram matrizes muito distintas que, conseqüentemente, geram topologias discordantes, indicando que todos menos um, ou todos, geram alinhamentos equivocados;
- As topologias resultantes dos diferentes programas de alinhamento concordam em muitos poucos casos com as relações filogenéticas geradas a partir de dados morfológicos;

- As limitações de amostragem podem ter um efeito significativo na qualidade dos resultados, sendo necessária uma análise com aumento dos táxons terminais, tanto dentro das subordens como dentro das famílias de maneira a otimizar os caracteres para o plano básico de cada clado e, possivelmente, aumentar a qualidade do resultado final;
- A monofilia de Diptera é recuperada somente para análises com os genes nucleares 18S e 28S;
- A infraordem Psychodomorpha (*sensu* Wood & Borkent, 1989) não foi recuperada em nenhuma das análises;
- O clado Neodiptera ( Bibionomorpha + Brachycera) é encontrado em algumas análises, sendo um dos poucos pontos de concordância com as análise baseadas em dados morfológicos;
- A infraordem Psychodomorpha resulta como parafilética em todas as análises se considerarmos Trichoceridae e Scatopsidae como parte dessa infraordem;
- Dois clados apareceram mais freqüentemente na base de Diptera - Culicomorpha e Psychodidae - sendo que o clado formado por Culicomorpha + Ptychopteromorpha é hipotetizado em uma das filogenias com dados morfológicos como grupo irmão do restante de Diptera. A presença de Psychodidae nessa posição é pouco provável. Ainda assim, a maior quantidade de dados e de hipóteses aponta para Tipulomorpha como grupo mais basal de Diptera, situação em que nenhuma das análises moleculares chegou à essa topologia.



## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- Amorim, D.S. 1993 (1992). A phylogenetic analysis of the basal groups of Bibionomorpha, with a critical examination of the wing vein homology. **Revista Brasileira de Biologia** 52 (3): 379-399.
- Amorim, D.S. 1994. A new suprageneric classification of the Scatopsidae (Diptera: Psychodomorpha). **Iheringia, Zoologia** 77: 107-112.
- Amorim, D.S. & C.J.B. de Carvalho. 1992. Book review: McAlpine, J.F. (ed.) & D.M. Wood (coord.) 1989. Manual of Nearctic Diptera, volume 3. **Revista Brasileira de Entomologia** 36 (4):899-904.
- Amorim, D. S.; Bravo, F. & Collucci, E. 1996. Thoracic sclerites: additional information concerning the basal evolution of Diptera (Insecta). **XX International Congress of Entomology, Proceedings**, Abstract 01-250, p. 64. Florença, Itália.
- Amorim, D. S. & Silva, V. C. 2002. How far advanced was Diptera evolution in the Pangea? **Annales de la Societe Entomologique de France** 38 (1-2): 177-200. (Procurar por: Souza, D. A. & Silva, V. C.).
- Amorim, D.S. & Grimaldi, D. 2006. Valseguyidae, a new family of Diptera in the Scatopsoidea, with a new genus in Cretaceous amber from Myanmar. **Systematic Entomology** 31(3) (on-line, impresso no número de agosto de 2006).
- Amorim, D.S. & Yeates, D. 2006. Pesky gnats: ridding dipteran classification of the Nematocera. **Studia Dipterologica** 13 (a aparecer no número de agosto de 2006).
- Amorim, D.S. & Rindal, E. Submetido. A phylogenetic study of the Mycetophiliformia, with creation of the subfamilies Heterotrichinae, Ohakuneinae, and Chiletrichinae for the Rangomaramidae (Diptera, Bibionomorpha). **Zoological Journal of the Linnean Society**
- Baker, R. H.; Wilkinson, G. S. & DeSalle, R. 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). **Systematic Biology** 50 (1): 87-105.
- Barrio, E.; Latorre, A. & Moya, A. 1994. Phylogeny of the *Drosophila obscura* species group deduced from mitochondrial DNA sequences. **Journal of Molecular Evolution** 39 (5): 478-88.
- Barrio, E.; Latorre, A.; Moya, A. & Ayala, F. J. 1992. Phylogenetic reconstruction of the *Drosophila obscura* group, on the basis of mitochondrial DNA. **Molecular Biology and Evolution** 9 (4): 621-35.
- Beati, L; Cáceres, A.G.; Lee, J.A. & Munstermann, L.E. 2004. Systematic relationships among *Lutzomyia* sand flies (Diptera: Psychodidae) of Peru and Colombia based on the analysis of

- 12S and 28S ribosomal DNA sequences. **International Journal for Parasitology** **34**: 225-234.
- Beebe, N.W.; Cooper, R.D.; Morrison, D.A. & Ellis, J.T. 2000. A phylogenetic study of the *Anopheles punctulatus* group of malaria vectors comparing rDNA sequence alignments derived from the mitochondrial and nuclear small ribosomal subunits. **Molecular Phylogenetics and Evolution** **17** (3): 430-436.
- Bernasconi, M. V.; Valsangiacomo, C.; Piffaretti, J. C. & Ward, P. I. 2000a. Phylogenetic relationships among Muscoidea (Diptera: Calyptratae) based on mitochondrial DNA sequences. **Insect Molecular Biology** **9** (1): 67-74.
- Bernasconi, M. V.; Pawlowski, J.; Valsangiacomo, C.; Piffaretti, J. C. & Ward, P. I. 2000b. Phylogeny of the Scathophagidae (Diptera, Calyptratae) based on mitochondrial DNA sequences. **Molecular Phylogenetics and Evolution** **16** (2): 308-15.
- Besansky, N. J.; Powell, J. R.; Caccone, A.; Hamm, D. M.; Scott, J. A. & Collins, F. H. 1994. Molecular phylogeny of the *Anopheles gambiae* complex suggests genetic introgression between principal malaria vectors. **Proceedings of the National Academy of Sciences of the United States of America** **91** (15): 6885-8.
- Blaschke-Berthold, U. 1994. Anatomie und Phylogenie der Bibionomorpha (Insecta, Diptera). **Bonner Zoologische Monographien** **34**: 1-206.
- Bridge, P.D.; Roberts, P.J.; Spooner, B.M. & Panchal, G. 2003. On the unreliability of published DNA sequences. **New Phytologist** **160**: 43-48.
- Caccone, A.; Moriyama, E. N.; Gleason, J. M.; Nigro, L. & Powell, J. R. 1996. A molecular phylogeny for the *Drosophila melanogaster* subgroup and the problem of polymorphism data. **Molecular Biology and Evolution** **13** (9): 1224-32.
- Carmean, D. & Crespi, J. B. 1995. Do long branches attract flies? **Nature** **373**: 666.
- Chen, B.; Harbach, R. E. & Butlin, R. K. 2002. Molecular and morphological studies on the *Anopheles minimus* group of mosquitoes in southern China: taxonomic review, distribution and malaria vector status. **Medical and Veterinary Entomology** **16** (3): 253-65.
- Colless, D.H. & McAlpine, D.K. 1991. Diptera (Flies). In: **The insects of Australia**. Melbourne University Press, CSIRO. 717-786.
- Collins, K.P. & Wiegmann, B.M. 2002a. Phylogenetic relationships and placement of the Empidoidea (Diptera: Brachycera) based on 28S rDNA and EF-1a sequences. **Insect Systematics and Evolution** **33**: 421-444.
- Collins, K.P. & Wiegmann, B.M. 2002b. Phylogenetic relationships of the Cyclorrhapha (Diptera: Brachycera) based on 28S rDNA. **Insect Systematics and Evolution** **33**: 445-456.

- Cook, C.E.; Austin, J.J. & Disney, R.H.L. 2004. A mitochondrial 12S and 16S rRNA phylogeny of critical genera of Phoridae (Diptera) and related families of Aschiza. **Zootaxa** **593**: 1-11.
- Courtney, G.W. 1990. Cuticular morphology of larval mountain midges (Diptera: Deuterophlebiidae): implications for the phylogenetic relationships of Nematocera. **Canadian Journal of Zoology** **68**: 556-578.
- Courtney, G.W. 1991. Phylogenetic analysis of the Blephariceromorpha, with special reference to mountain midges (Diptera: Deuterophlebiidae). **Systematic Entomology** **16**: 137-172.
- Damgaard, J.; Andersen, N.M. & Meier, R. 2005. Combining molecular and morphological analyses of water strider phylogeny (Hemiptera-Heteroptera, Gerromorpha): effects of alignment and taxon sampling. **Systematic Entomology** **30**: 289-309.
- Depaquit, J.; Ferte, H.; Leger, N.; Killick-Kendrick, R.; Rioux, J. A.; Killick-Kendrick, M.; Hanafi, H. A. & Gobert, S. 2000. Molecular systematics of the phlebotomine sandflies of the subgenus *Paraphlebotomus* (Diptera, Psychodidae, *Phlebotomus*) based on ITS2 rDNA sequences. Hypotheses of dispersion and speciation. **Insect Molecular Biology** **9** (3): 293-300.
- DeSalle, R. 1992. The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. **Molecular Phylogenetics and Evolution** **1** (1): 31-40.
- Di Muccio, T.; Marinucci, M.; Frusteri, L.; Maroli, M.; Pesson, B. & Gramiccia, M. 2000. Phylogenetic analysis of *Phlebotomus* species belonging to the subgenus *Larroussius* (Diptera, Psychodidae) by ITS2 rDNA sequences. **Insect Biochemistry and Molecular Biology** **30** (5): 387-93.
- Dujardin, J.P.; Ponto, F.L. & Martinez, E. 1999. Quantitative phenetics and taxonomy of some phlebotomine taxa. **Memórias do Instituto Oswaldo Cruz** **94**(6): 735-741.
- Durando, C. M.; Baker, R. H.; Etges, W. J.; Heed, W. B.; Wasserman, M. & DeSalle, R. 2000. Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. **Molecular Phylogenetics and Evolution** **16** (2): 296-307.
- Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. **Systematic Zoology** **27**: 401-410.
- Felsenstein, J. 1982. How can we infer geography and history from gene frequencies? **Journal of Theoretical Biology** **96**: 9-20.
- Foley, D. H.; Bryan, J. H.; Yeates, D. & Saul, A. 1998. Evolution and systematics of *Anopheles*: insights from a molecular phylogeny of Australasian mosquitoes. **Molecular Phylogenetics and Evolution** **9** (2): 262-75.
- Friedrich, M. & Tautz, D. 1997a. Evolution and phylogeny of the Diptera: a molecular phylogenetic analysis using 28S rDNA sequences. **Systematic Biology** **46** (4): 674-698.

- Friedrich, M. & Tautz, D. 1997. An episodic change of rDNA nucleotide substitution rate has occurred during the emergence of the insect order Diptera. **Molecular Biology and Evolution** **14** (6): 644-653.
- Gimeno, C.; Belshaw, R. & Quicke, D.L.J. 1997. Phylogenetic relationships of the Alysiniinae?Opiinae (Hymenoptera: Braconidae) and the utility of cytochrome b, 16S and 28S D2 rRNA. **Insect Molecular Biology** **6** (3): 273-284.
- Gimnig, J. E. & Eldridge, B. F. 1999. Genetic and morphological characterization of the *Aedes* (*Ochlerotatus*) *dorsalis* (Diptera: Culicidae) group in North America. **Journal of Medical Entomology** **36** (6): 685-94.
- Giribet, G. & Wheeler, W.C. 1999. On gaps. **Molecular Phylogenetics and Evolution** **13** (1): 132-143.
- Giribet, G.; Distel, D.L.; Polz, M.; Sterrer, W. & Wheller, W.C. 2000. Triploblastic relationships with emphasis on the acoelomates and the position of Gnathostomulida, Cyclophora, Plathelminthes, and Chaetognatha: A combined approach of 18S rDNA sequences and morphology. **Systematic Biology** **49**:539–562.
- Giribet, G.; Edgecombe, G.D. & Wheller, W.C. 2001. Arthropod phylogeny based on eight molecular loci and morphology. **Nature** **413**: 157–161.
- Giribet, G.; Edgecombe, G.D.; Wheeler, W.C. & Babbit, C. 2002. Phylogeny of the Arachnida and Opiliones: a combined approach using morphological and molecular sequence data. **Cladistics** **18**: 5–70.
- Graham, R.L. & Foulds, L. R. 1982. Unlikelihood that minimal phylogenies for a realistic biological study can be constructed in reasonable computational time. **Mathematical Biosciences** **60**: 133-142.
- Griffiths, G.C.D. 1990. Review of Manual of Nearctic Diptera. Volume 3. J.F. McAlpine and D.M. Wood (editors). **Quaestiones Entomologicae** **26**: 117-130.
- Griffiths, G.C.D. 1994. Relationships among the major subgroups of Brachycera (Diptera): a critical review. **Canadian Entomologist** **126**: 861-880.
- Griffiths, G.C.D. 1996. Review of papers on the male genitalia of Diptera by D.M. Wood and associates. **Studia Dipterologica** **3**: 107-123.
- Guryev, V.; Makarevitch, I.; Blinov, A. & Martin, J. 2001. Phylogeny of the genus *Chironomus* (Diptera) inferred from DNA sequences of mitochondrial cytochrome b and cytochrome oxidase I. **Molecular Phylogenetics and Evolution** **19** (1): 9-21.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series** **41**, 95–98.

- Han, H. Y. & McPherson, B. A. 1997. Molecular phylogenetic study of Tephritidae (Insecta: Diptera) using partial sequences of the mitochondrial 16S ribosomal DNA. **Molecular Phylogenetics and Evolution** 7 (1): 17-32.
- Hayward, D.C.; Trueman, J.W.H.; Bastiani, M.J. & Ball, E.E. 2005. The structure of the USP/RXR of *Xenos pecki* indicates that Strepsiptera are not closely related to Diptera. **Development Genes and Evolution** 215: 213-219.
- Hein, J. 1989. A new method that simultaneously aligns and reconstructs ancestral sequences for any number of homologous sequences, when a phylogeny is given. **Molecular Biology and Evolution** 6: 649-668.
- Hein, J. 1990. Unified approach to alignment and phylogenetics. **Methods in Enzymology** 183: 626-644.
- Hennig, W. 1966. **Phylogenetic systematic**. Univ. Illinois Press. Urbana. 263 p.
- Hennig, W. 1968. Kritische Bemerkungen über den Bau der Flügelwurzel bei den Diptern und die Frage nach der Monophylie der Nematocera. **Stuttgarter Beiträge zur Naturkunde** 195: 1-23.
- Hennig, W. 1973. Diptera (Zweiflügler). pages 1-337, *In* Helmcke *et al.* (eds): **Handbuch der Zoologie. Eine Naturgeschichte der Stämme des Tierreiches**. IV. Band: Arthropoda - 2. Hälfte: Insecta. 2. Teil: Spezielles. 31. Berlin - New York.
- Hennig, W. 1981. **Insect phylogeny**. New York: Wiley.
- Hertwig, S.; de Sá, R.O. & Haas, A. 2004. Phylogenetic signal and the utility of 12S and 16S mtDNA in frog phylogeny. **Journal of Zoological Systematics and Evolutionary Research** 42: 2-18.
- Higgins, D.G.; Bleasby, A.J. & Fuchs, R. 1992. CLUSTAL V: Improved software for multiple sequence alignment. **Computer Applications in the Biosciences** 8: 189-191.
- Higgins, D. G. & Sharp, P. M. 1988. CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. **Gene** 73: 237-244.
- Higgins, D.G.; Thompson, J.D. & Gibson, T.J. 1996. Using CLUSTAL for multiple sequence alignments. **Methods in Enzymology** 266: 383-402.
- Hovenkamp, P. 2004. Review of: T.N.T.-Tree Analysis Using New Technology. Version 1.0, by P. Goloboff, J.S. Farris and K. Nixon. Available from authors and from <http://www.zmuc.dk/public/phylogeny>. **Cladistics** 20: 378-383.
- Huelsenbeck, J. P. 1997. Is the Felsenstein zone a fly trap? **Systematic Biology** 46: 69-74.
- 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. **Molecular Phylogenetics and Evolution** 9 (3): 470-480.

- Ishikawa, E. A.; Ready, P. D.; de Souza, A. A.; Day, J. C.; Rangel, E. F.; Davies, C. R. & Shaw, J. J. 1999. A mitochondrial DNA phylogeny indicates close relationships between populations of *Lutzomyia whitmani* (Diptera: Psychodidae, Phlebotominae) from the rain-forest regions of Amazonia and northeast Brazil. **Memórias do Instituto Oswaldo Cruz** **94** (3): 339-45.
- Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. 1998. Multiple sequence alignment with Clustal X. **Trends in Biochemical Sciences** **23**: 403–405.
- Kambhampati, S. & Charlton, R.E. 1999. Phylogenetic relationship among *Libellula*, *Ladona* and *Plathemis* (Odonata: Libellulidae) based on DNA sequence of mitochondrial 16S rRNA gene. **Systematic Entomology** **24**: 37-49.
- Kambhampati, S.; Volkll, W. & Mackauer, M. 2000. Phylogenetic relationships among genera of Aphidiinae (Hymenoptera: Braconidae) based on DNA sequence of the mitochondrial 16S rRNA gene. **Systematic Entomology** **25**: 437-445.
- Kluge, A.G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidea, Serpentes). **Systematic Zoology** **38**: 7-25.
- Kluge, A.G. & Wolf, A.J. 1993. Cladistics: what's in a word? **Cladistics** **9**: 183-199.
- Kristensen, N. P. 1975. The phylogeny of hexapod "orders". A critical review of recent accounts. **Zeitschrift für Zoologische Systematik und Evolutionsforschung** **13**: 1-44.
- Krzeminski, W. 1992. Triassic and Lower Jurassic stage of Diptera Evolution. **Mitteilungen Der Schweizerischen Entomologischen Gesellschaft** **65**: 39-59.
- Krzywinski, J. & Besansky, N. J. 2003. Molecular systematics of *Anopheles*: from subgenera to subpopulations. **Annual Review of Entomology** **48**: 111-39.
- Kukalová-Peck, J. 1991. Fossil history and the evolution of hexapod structures. *In*: CSIRO. **The insects of Australia**. Melbourne: Melbourne Univ. Press.
- Kurahashi, H. & Kano, R. 1984. Phylogeny and geographical distribution of the genus *Boettcherisca* Rohdendorf (Diptera: Sarcophagidae). **Japanese Journal of Medical Science and Biology** **37** (1): 27-34.
- Kwiatowski, J. & Ayala, F. J. 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. **Molecular Phylogenetics and Evolution** **13** (2): 319-28.
- Mazzarolo, L.A. & Amorim, D.S. 2000. *Cratomyia macrorrhyncha*, a Lower Cretaceous brachyceran fossil from the Snatana Formation, Brazil, representing a new species, genus and family of the Stratiomyomorpha (Diptera). **Insect Systematics & Evolution** **31**: 91-102.

- McAlpine, J.F.; Peterson, B.V.; Shewell, G.E.; Teskey, H.J.; *et al.* 1981. **Manual of Nearctic Diptera**, Vol 1. Ottawa, Can. Res. Branch Agric. Can.
- McAlpine, J.F. & Wood, D.V. (Eds) 1989. **Manual of Nearctic Diptera**, Vol 3. Ottawa, Can. Res. Branch Agric. Can.
- McAlpine, J. F. 1989. Phylogeny and classification of the Muscomorpha. *In*: McAlpine, J. F. & Wood, D. M. Eds. **Manual of Nearctic Diptera**, Vol 3. Ottawa, Can: Res. Branch Agric. Can.
- Meier, R. & Wiegmann, B. M. 2002. A phylogenetic analysis of Coelopidae (Diptera) based on morphological and DNA sequence data. **Molecular Phylogenetics and Evolution** **25** (3): 393-407.
- Meigen, J.W. 1800. **Nouvelle classification des mouches à deux ailes (Diptera L.) d'après un plan tout nouveau**. Paris: Perronneau 40 pp.
- Meigen, J.W. 1803. Versuch einer neuen Gattungs Eintheilung der europäischen zweiflügligen Insekten. **Magazin für Insektenkunde (Illiger)** **2**: 259-281.
- Michelsen, V. 1996. Neodiptera: new insights into the adult morphology and higher level phylogeny of Diptera (Insecta). **Zoological Journal of the Linnean Society** **117**: 71-102.
- Mickevich, M. F. & Farris, J. S. 1981. The implications of congruence in *Menidia*. **Systematic Zoology** **30**: 351-370.
- Miller, B. R.; Crabtree, M. B. & Savage, H. M. 1997. Phylogenetic relationships of the Culicomorpha inferred from 18S and 5.8S ribosomal DNA sequences (Diptera: Nematocera). **Insect Molecular Biology** **6** (2): 105-14.
- Morrison, D.A. & Ellis, J.T. 1997. Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of Apicomplexa. **Molecular Biology and Evolution** **14** (4): 428-441.
- Moulton, J.K. 2000. Molecular sequence data resolves basal divergences within Simuliidae (Diptera). **Systematic Entomology** **25**:95-113.
- Muraji, M. & Nakahara, S. 2001. Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera, Tephritidae), based on the mitochondrial rDNA sequences. **Insect Molecular Biology** **10** (6): 549-59.
- Nikolaidis, N. & Scouras, Z. G. 1996. The *Drosophila montium* subgroup species. Phylogenetic relationships based on mitochondrial DNA analysis. **Genome** **39** (5): 874-83.
- Nirmala, X.; Hypsa, V. & Zurovec, M. 2001. Molecular phylogeny of Calyptratae (Diptera: Brachycera): the evolution of 18S and 16S ribosomal rDNAs in higher dipterans and their use in phylogenetic inference. **Insect Molecular Biology** **10** (5): 475-85.

- Nixon, C. K. & Carpenter, J. M. 1993. On outgroups. **Cladistics** **9**: 413-426.
- Nurminsky, D. I.; Moriyama, E. N.; Lozovskaya, E. R. & Hartl, D. L. 1996. Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplications of the alcohol dehydrogenase gene. **Molecular Biology and Evolution** **13** (1): 132-49.
- Ogden, T.H. & Whiting, M.F. 2003. The problem with "the Paleoptera problem:" sense and sensitivity. **Cladistics** **19**:432-442.
- O'Grady, P. M. 1999. Reevaluation of phylogeny in the *Drosophila obscura* species group based on combined analysis of nucleotide sequences. **Molecular Phylogenetics and Evolution** **12** (2): 124-39.
- O'Grady, P. M.; Clark, J. B. & Kidwell, M. G. 1998. Phylogeny of the *Drosophila saltans* species group based on combined analysis of nuclear and mitochondrial DNA sequences. **Molecular Biology and Evolution** **15** (6): 656-64.
- O'Grady, P. M. & Kidwell, M. G. 2002. Phylogeny of the subgenus *sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. **Molecular Phylogenetics and Evolution** **22** (3): 442-53.
- Oosterbroek, P. & Theowald, B. 1991. Phylogeny of the Tipuloidea based on characters of larvae and pupae (Diptera, Nematocera). **Tijdschrift voor Entomologie** **134**:211-267.
- Oosterbroek, P. & Courtney, G. 1995. Phylogeny of the nematoceros families of Diptera (Insecta). **Zoological Journal of the Linnean Society** **115**: 267-311.
- Pelandakis, M.; Higgins, D. G. & Solignac, M. 1991. Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. **Genetica** **84** (2): 87-94.
- Phillips, A.; Janeis, D. & Wheeler, W. 2000. Multiple sequence alignment in phylogenetic analysis. **Molecular Phylogenetics and Evolution** **16** (3): 317-330.
- Piano, F.; Craddock, E. M. & Kambysellis, M. P. 1997. Phylogeny of the island populations of the Hawaiian *Drosophila grimshawi* complex: evidence from combined data. **Molecular Phylogenetics and Evolution** **7** (2): 173-84.
- Pollock, D.D.; Zwickl, D.J. McGuire, J.A. & Hillis, D.M. 2002. Increased taxon sampling is advantageous for phylogenetic inference. **Systematic Biology** **51**: 664-671.
- Porter, C. H. & Collins, F. H. 1996. Phylogeny of nearctic members of the *Anopheles maculipennis* species group derived from the D2 variable region of 28S ribosomal RNA. **Molecular Phylogenetics and Evolution** **6** (2): 178-88.
- Pruess, K. P.; Adams, B. J.; Parsons, T. J.; Zhu, X. & Powers, T. O. 2000. Utility of the mitochondrial cytochrome oxidase II gene for resolving relationships among black flies (Diptera: Simuliidae). **Molecular Phylogenetics and Evolution** **16** (2): 286-95.



- Ready, P. D.; de Souza, A. A.; Rebelo, J. M.; Day, J. C.; Silveira, F. T.; Campbell-Lendrum, D.; Davies, C. R. & Costa, J. M. 1998. Phylogenetic species and domesticity of *Lutzomyia whitmani* at the southeast boundary of Amazonian Brazil. **Transactions of the Royal Society of Tropical Medicine and Hygiene** **92** (2): 159-60.
- Rokas, A.; Kathirithamby, J. & Holland, P.W.H. 1999. Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. **Insect Molecular Biology** **8** (4): 527-530.
- Rosenberg, M. S. & Kumar, S. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. **Proceedings of the National Academy of Sciences USA** **908**: 10751-10756.
- Russo, C. A. M. 2001. Como escolher genes para problemas filogenéticos específicos. *In*: Mاتيoli, S. R. (Ed.). **Biologia molecular e evolução**. Editora Holos, Ribeirão Preto, São Paulo, pp. 130-136.
- Saether, O.A. 2000. Phylogeny of Culicomorpha. **Systematic Entomology** **25**: 223-234.
- Sallum, M. A.; Bergo, E. S.; Flores, D. C. & Forattini, O. P. 2002. Systematic studies on *Anopheles galvaoi* Causey, Deane & Deane from the subgenus *Nyssorhynchus* Blanchard (Diptera: Culicidae). **Memórias do Instituto Oswaldo Cruz** **97** (8): 1177-89.
- Sanderson, M. J. & Shaffer, H. B. 2002. Troubleshooting molecular phylogenetic analyses. **Annual Review of Ecology and Systematics** **33**: 49-72.
- Savage, J.; Wheeler, T.A. & Wiegmann, B.M. 2004. Phylogenetic analysis of the genus *Thricops* Rondani (Diptera: Muscidae) based on molecular and morphological characters. **Systematic Entomology** **29** (3): 395-414.
- Scheffer, S. J. & Wiegmann, B. M. 2000. Molecular phylogenetics of the holly leafminers (Diptera: Agromyzidae: *Phytomyza*): species limits, speciation, and dietary specialization. **Molecular Phylogenetics and Evolution** **17** (2): 244-55.
- Schuehli, G.S; Carvalho, C.J.B. & Wiegmann, B.M. 2004. Regarding the taxonomic status of *Ophyra* Robineau-Desvoid (Diptera: Mucidae): a molecular approach. **Zootaxa** **712**: 1-12.
- Simon, C.; Frati, F.; Beckenbach, A.; Crespi, B.; Liu, H. & Flook, P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gen sequences and a compilation of conserved polymerase chain reaction primers. **Annals of the Entomological Society of America** **87**: 651-686.
- Sinclair B.J. 1992. A phylogenetic interpretation of the Brachycera (Diptera) based on the larval mandible and associated mouthpart structures. **Systematic Entomology** **17**: 233-252.

- Sinclair B.J. & Cumming, J.M. 2006. The morphology, higher-level phylogeny and classification of the Empidoidea (Diptera). **Zootaxa** **1180**: 172 pp.
- Smith, J. J. & Bush, G. L. 1997. Phylogeny of the genus *Rhagoletis* (Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase II. **Molecular Phylogenetics and Evolution** **7** (1): 33-43.
- Smith, P. T.; Kambhampati, S. & Armstrong, K. A. 2003. Phylogenetic relationships among *Bactrocera* species (Diptera: Tephritidae) inferred from mitochondrial DNA sequences. **Molecular Phylogenetics and Evolution** **26** (1): 8-17.
- Spicer, G. S. & Pitnick, S. 1996. Molecular systematics of the *Drosophila hydei* subgroup as inferred from mitochondrial DNA sequences. **Journal of Molecular Evolution** **43** (3): 281-6.
- Stahls, G. & Nyblom, K. 2000. Phylogenetic analysis of the genus *Cheilosia* (Diptera, Syrphidae) using mitochondrial COI sequence data. **Molecular Phylogenetics and Evolution** **15** (2): 235-41.
- Suzuki, Y. M.; Kitagawa, O. & Wakahama, K. I. 1990. Chromosomal analysis and phylogenetic relationships in the *Drosophila nasuta* subgroup. I. Phylogenetic relationships within the *Drosophila sulfurigaster* species complex. **Genetica** **80** (1): 53-66.
- Svenson, G.J. & Whiting, M.F. 2004. Phylogeny of Mantodea based on molecular data: evolution of a charismatic predator. **Systematic Entomology** **29**: 359–370.
- Swofford, D. L. 2000. PAUP\* 4.0: Phylogenetic analysis Using Parsimony (\*and Other Methods). Sinauer, Sunderland, MA
- Swofford, D.L. 2002. PAUP\*: Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.
- Terry, M. D. & Whiting, M.F. 2005. Comparison of two alignment techniques within a single data set: POY v. CLUSTAL. **Cladistics** **21**(3): 272-281.
- Thomas, R. H. & Hunt, J. A. 1993. Phylogenetic relationships in *Drosophila*: a conflict between molecular and morphological data. **Molecular Biology and Evolution** **10** (2): 362-74.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research** **22**: 4673–4680.
- 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** **25** (24): 4876-4821.
- Wheeler, W.C. 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. **Systematic Biology** **44** (3): 321-331.

- Wheeler, W. 1996. Optimization alignment: the end of multiple sequence alignment in phylogenetics? **Cladistics** **12**:1-9.
- Wheeler, W.C. 1999. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. **Systematic Biology** **44** (3): 321-331.
- Wheeler, W. 2001. Homology and the optimization of DNA sequence data. **Cladistics** **17**: S3-S11.
- Wheeler, W. 2003. Search-based optimization. **Cladistics** **19**: 348-355.
- Wheeler, W.C. & Gladstein, D.S. 1994. MALIGN: a multiple sequence alignment program. **Journal of Heredity** **85**: 417-418.
- Wheeler, W. C. & Hayashi, C. Y. 1998. The phylogeny of extant Chelicerate orders. **Cladistics** **14**: 173-192.
- Wheeler, W.C.; Gladstein, D. & Laet, J.D. 2003. POY, Phylogeny Reconstruction via Optimization of DNA and other Data. Version 3.0.11 . American Museum of Natural History, New York. Programa e documentação disponíveis em <ftp://ftp.amnh.org/pub/molecular/poy/>
- Whiting, M.F. 1998. Phylogenetic position of the Strepsiptera: review of molecular and morphological evidence. **International Journal of Insect Morphology and Embryology** **27**: 53–60.
- Whiting, M.F., 2001. Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. **Zoologica Scripta** **31**: 93–104.
- Whiting, M.F. 2002. Phylogeny of the holometabolous insect orders: molecular evidence. **Zoologica Scripta** **31**: 3-15.
- Whiting, M. F. & Wheeler, W. C. 1994. Insect homeotic transformation. **Nature** **368**: 696.
- Whiting, M.F.; Carpenter, J.C.; Wheeler, Q.D. & Wheeler, W.C. 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. **Systematic Biology** **46**(1):1-68.
- 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.; Tsaur, S.C.; Webb, D.W.; Yeates, D.K. & Cassel, B.K. 2000. Monophyly and relationships of the Tabanomorpha (Diptera: Brachycera) based on 28S ribosomal gene sequences. **Annals of the Entomological Society of America** **93**(5): 1031-1038.
- Wiens, J. J. 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? **Systematic Biology** **47**:625-640.

- Wood, D.M. 1991. Homology and phylogenetic implications of male genitalia in Diptera. The ground plan. In: L. Weisman, I. Orzogh & A.C. Pont (Eds.): **Proceedings of the 2nd International Congress on Dipterology**, The Hague, pp. 255-284.
- Wood, D. M. & Borkent, A. 1982. Description of the female of *Parasimulium crosskeyi* Peterson (Diptera: Simuliidae) and a discussion of the phylogenetic position of the genus. **Memoirs of the Entomological Society of Washington 10**: 193-210.
- Wood, D. M. & Borkent, A. 1989. Phylogeny and classification of the Nematocera. *In*: McAlpine, J. F. & Wood, D. M. Eds. **Manual of Nearctic Diptera**, Vol 3. Ottawa, Can: Res. Branch Agric. Can.
- Woodley, N. E. 1989. Phylogeny and classification of the "Orthorrhaphous" Brachycera. *In*: McAlpine, J. F. & Wood, D. M. Eds. **Manual of Nearctic Diptera**, Vol 3. Ottawa, Can: Res. Branch Agric. Can.
- Yeates, D.K. 1994. The cladistics and classification of the Bombyliidae (Diptera: Asiloidea). **Bulletin of the American Museum of Natural History, 219**: 1-191. (<http://hdl.handle.net/2246/826>)
- Yaetes, D. K. & Wiegmann, B. M. 1999. Congruence and controversy: toward a higher-level phylogeny of Diptera. **Annual Review of Entomology 44**: 397-428.
- Zwickl, D. & Hillis, D.M. 2002. Increased taxon sampling greatly reduces phylogenetic error. **Systematic Biology 51** (4): 588-598.

## Anexo I

Táxons terminais utilizados na primeira análise do gene mitocondrial ribossomal 12S.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Trirhabda virgata</i>	Coleoptera			AY507343
<i>Trirhabda pilosa</i>	Coleoptera			AY507338
<i>Nestinus bimaculatus</i>	Coleoptera			AY507271
<i>Pyrrhalta luteola</i>	Coleoptera			AY507268
<i>Percus lineatus</i>	Coleoptera			AY334359
<i>Achenomorphus corticinus</i>	Coleoptera			AF021066
<i>Platyphora amabilis</i>	Coleoptera			AY055581
<i>Oreina intricata</i>	Coleoptera			AF097071
<i>Carabus splendens</i>	Coleoptera			AF190030
<i>Pterostichus niger</i>	Coleoptera			AF190029
<i>Percus villai</i>	Coleoptera			AF190026
<i>Abax parallelepipedus inferior</i>	Coleoptera			AF190024
<i>Dyscia penulataria</i>	Lepidoptera			AJ505770
<i>Antheraea yamamai</i>	Lepidoptera			AY037835
<i>Bombix mandarina</i>	Lepidoptera			AY037828
<i>Micromus</i> sp.	Neuroptera			AY139924
<i>Ctenocephalides felis</i>	Siphonaptera			U73741
<i>Baldratia oculata</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495754
<i>Baldratia salicorniae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495744
<i>Baldratia suadae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495745
<i>Careopalpis</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495753
<i>Izeniola</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495752
<i>Izeniola</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495751
<i>Izeniola</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495747
<i>Izeniola obesula</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495742
<i>Mayetiola destructor</i>	Diptera	Bibionomorpha	Cecidomyiidae	AY460211
<i>Orselia oryzae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AF306555
<i>Orselia oryzae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AF306554
<i>Orselia oryzae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AF306553
<i>Stefaniella trinacriacae</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495756
<i>Stefaniola brevis</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495743
<i>Stefaniola</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495755
<i>Stefaniola</i> sp.	Diptera	Bibionomorpha	Cecidomyiidae	AJ495748
<i>Stefaniola siliqua</i>	Diptera	Bibionomorpha	Cecidomyiidae	AJ495749
<i>Lutzomyia serrana</i>	Diptera	Psychodomorpha	Psychodidae	AY352699
<i>Lutzomyia tejadai</i>	Diptera	Psychodomorpha	Psychodidae	AY352701
<i>Lutzomyia tuberculata</i>	Diptera	Psychodomorpha	Psychodidae	AY352703
<i>Lutzomyia verrucarum</i>	Diptera	Psychodomorpha	Psychodidae	AY352705
<i>Lutzomyia yuilli yuilli</i>	Diptera	Psychodomorpha	Psychodidae	AY352707
<i>Lutzomyia</i> sp.	Diptera	Psychodomorpha	Psychodidae	AY352708
<i>Warileyia euniceae</i>	Diptera	Psychodomorpha	Psychodidae	AY352709
<i>Warileyia phlebotomanica</i>	Diptera	Psychodomorpha	Psychodidae	AY352710
<i>Metricnemus</i> sp.	Diptera	Culicomorpha	Chironomidae	AF304344
<i>Anopheles hilli</i>	Diptera	Culicomorpha	Culicidae	L02382
<i>Austrosimulium bancrofti</i>	Diptera	Culicomorpha	Simuliidae	L02399
<i>Austrosimulium bancrofti</i>	Diptera	Culicomorpha	Simuliidae	L02398
<i>Austrosimulium bancrofti</i>	Diptera	Culicomorpha	Simuliidae	L02390
<i>Austrosimulium bancrofti</i>	Diptera	Culicomorpha	Simuliidae	L02386
<i>Austrosimulium pestilens</i>	Diptera	Culicomorpha	Simuliidae	L02385
<i>Scaptia</i> sp.	Diptera	Tabanomorpha	Tabanidae	L02393

<i>Hilara maura</i>	Diptera	Asilomorpha	Empididae	AF126318
<i>Lonchoptera lutea</i>	Diptera	Aschiza	Lonchopteridae	AF304342
<i>Lonchoptera lutea</i>	Diptera	Aschiza	Lonchopteridae	AF126310
<i>Anerina thoracica</i>	Diptera	Aschiza	Phoridae	AF126317

## Anexo I

### Continuação

<i>Chonocephalus beymonsi</i>	Diptera	Aschiza	Phoridae	AF126324
<i>Chonocephalus beymonsi</i>	Diptera	Aschiza	Phoridae	AF126313
<i>Clitelloxenia audreyae</i>	Diptera	Aschiza	Phoridae	AF126311
<i>Conicera similis</i>	Diptera	Aschiza	Phoridae	AF126309
<i>Diploneura mortimeri</i>	Diptera	Aschiza	Phoridae	AF126316
<i>Diploneura nitidula</i>	Diptera	Aschiza	Phoridae	AF126298
<i>Dobrniphora trigonae</i>	Diptera	Aschiza	Phoridae	AF126315
<i>Megaselia scalaris</i>	Diptera	Aschiza	Phoridae	AF126312
<i>Megaselia aequalis</i>	Diptera	Aschiza	Phoridae	AF126300
<i>Metopina pileata</i>	Diptera	Aschiza	Phoridae	AF126301
<i>Puliciphora borinquenensis</i>	Diptera	Aschiza	Phoridae	AF126323
<i>Phora atra</i>	Diptera	Aschiza	Phoridae	AF126304
<i>Spiniphora punctipennis</i>	Diptera	Aschiza	Phoridae	AF126314
<i>Termitophilomyia zimbrensi</i>	Diptera	Aschiza	Phoridae	AF126322
<i>Termitophilomyia zimbrensi</i>	Diptera	Aschiza	Phoridae	AF126321
<i>Thaumatoxena andreinii</i>	Diptera	Aschiza	Phoridae	AF126325
<i>Tripbleba distinguenda</i>	Diptera	Aschiza	Phoridae	AF126303
<i>Tripbleba nudipalpis</i>	Diptera	Aschiza	Phoridae	AF126302
<i>Tripbleba distinguenda</i>	Diptera	Aschiza	Phoridae	AF126299
<i>Protochybia modesta</i>	Diptera	Aschiza	Platyppezidae	AF126320
<i>Melanostoma</i> sp.	Diptera	Aschiza	Syrphidae	AF304343
<i>Melanostoma scalaris</i>	Diptera	Aschiza	Syrphidae	AF126308
<i>Agromyzidae</i> sp.	Diptera	Acalyptratae	Agromyzidae	AF304336
<i>Melanagromyza minimoides</i>	Diptera	Acalyptratae	Agromyzidae	AY573083
<i>Oscinella frit</i>	Diptera	Acalyptratae	Chloropidae	AF304346
<i>Limnobia quadrata</i>	Diptera	Acalyptratae	Ephydriidae	AF126305
<i>Hydrellia modesta</i>	Diptera	Acalyptratae	Ephydriidae	AF126306
<i>Rainieria antennae</i>	Diptera	Acalyptratae	Micropezidae	AY573098
<i>Sepsis cynipsea</i>	Diptera	Acalyptratae	Sepsidae	AF304349
<i>Sepsis fulgens</i>	Diptera	Acalyptratae	Sepsidae	AF304348
<i>Lotophila atra</i>	Diptera	Acalyptratae	Sphaeroceridae	AF304341
<i>Delia platura</i>	Diptera	Calypttratae	Anthomyiidae	AF304338
<i>Delia florilega</i>	Diptera	Calypttratae	Anthomyiidae	AF304337
<i>Lucilia cuprina</i>	Diptera	Calypttratae	Calliphoridae	L02391
<i>Australofannia</i> sp.	Diptera	Calypttratae	Fanniidae	L02401
<i>Musca domestica</i>	Diptera	Calypttratae	Muscidae	L02409
<i>Musca vetustissima</i>	Diptera	Calypttratae	Muscidae	L02400
<i>Scathophaga stercoraria</i>	Diptera	Calypttratae	Scathophagidae	AF304350
<i>Ceromya fergusonii</i>	Diptera	Calypttratae	Tachinidae	L02387
<i>Macropodexia</i> sp.	Diptera	Calypttratae	Tachinidae	L02384

## Anexo II

Táxons terminais utilizados na segunda análise do gene mitocondrial ribossomal 12S.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Athous</i> sp.	Coleoptera			AF232894
<i>Coccotrypes dactyliperda</i>	Coleoptera			AF444052
<i>Cyclorhpidium agnatum</i>	Coleoptera			AF249861
<i>Quercusia quercus</i>	Lepidoptera			AF232891
<i>Ideopsis similes</i>	Lepidoptera			AY256845
<i>Torticodes alternella</i>	Lepidoptera			AF232892
<i>Mantis religiosa</i>	Mantodea			U17792
<i>Bittacomorpha clavipes</i>	Diptera	Ptychopteromorpha	Ptychopteridae	AF049468
<i>Cardiocladius</i> sp.	Diptera	Culicomorpha	Chironomidae	AF049470
<i>Anopheles annulipes</i>	Diptera	Culicomorpha	Culicidae	AF121064
<i>Anopheles farauti</i>	Diptera	Culicomorpha	Culicidae	AF121071
<i>Anopheles farauti</i>	Diptera	Culicomorpha	Culicidae	AF121069
<i>Dixa</i> sp.	Diptera	Culicomorpha	Dixidae	AF049469
<i>Parasimulium crosskeyi</i>	Diptera	Culicomorpha	Simuliidae	AF049472
<i>Helodon onychodactylus</i>	Diptera	Culicomorpha	Simuliidae	AF049474
<i>Cnephia ornithophilina</i>	Diptera	Culicomorpha	Simuliidae	AF049477
<i>Ectemnia reclusa</i>	Diptera	Culicomorpha	Simuliidae	AF049479
<i>Thaumalea gillespieae</i>	Diptera	Culicomorpha	Thaumaleidae	AF049471
<i>Megaselia</i> sp.	Diptera	Aschiza	Phoridae	AF154729
<i>Megaselia</i> sp.	Diptera	Aschiza	Phoridae	AF154728
<i>Chalarus</i> sp.	Diptera	Aschiza	Pipunculidae	AF154686
<i>Chalarus</i> sp.	Diptera	Aschiza	Pipunculidae	AF154685
<i>Collinias</i> sp.	Diptera	Aschiza	Pipunculidae	AF154693
<i>Lindneromyia argentifascia</i>	Diptera	Aschiza	Platypezidae	AF154732
<i>Cyphipelta rufocyanea</i>	Diptera	Aschiza	Syrphidae	AF154695
<i>Melanostoma apicale</i>	Diptera	Aschiza	Syrphidae	AF154714
<i>Dideopsis aegrotata</i>	Diptera	Aschiza	Syrphidae	AF154697
<i>Neoconops longicornis</i>	Diptera	Acalyptratae	Conopidae	AF154735
<i>Neoconops longicornis</i>	Diptera	Acalyptratae	Conopidae	AF154734
<i>Teloglabus milleri</i>	Diptera	Acalyptratae	Diopsidae	AF304697
<i>Cyrtodiopsis currani</i>	Diptera	Acalyptratae	Diopsidae	AF304711
<i>Cyrtodiopsis quinqueguttata</i>	Diptera	Acalyptratae	Diopsidae	AF304713
<i>Mimegralla australica</i>	Diptera	Acalyptratae	Micropezidae	AF154718
<i>Delphinia picta</i>	Diptera	Acalyptratae	Otitidae	AY037484
<i>Zonosemata electa</i>	Diptera	Acalyptratae	Tephritidae	AY037505
<i>Bactrocera occipitales</i>	Diptera	Acalyptratae	Tephritidae	AY037343
<i>Adrama apicalis</i>	Diptera	Acalyptratae	Tephritidae	AY037498
<i>Chrysonya putoria</i>	Diptera	Calypttratae	Calliphoridae	NC_002697
<i>Mesembrina meridiana</i>	Diptera	Calypttratae	Muscidae	AF232899

### Anexo III

Táxons terminais utilizados na primeira análise do gene mitocondrial ribossomal 16S.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Timarcha</i> sp.	Coleoptera			AJ231160
<i>Sesamia nonagrioides</i>	Lepidoptera			AY320467
<i>Scirpophaga excerptalis</i>	Lepidoptera			AY320461
<i>Opogona glycyphaga</i>	Lepidoptera			AY320459
<i>Cosmopteryx</i> sp.	Lepidoptera			AY320442
<i>Mantodea</i> sp.	Mantodea			AY318882
<i>Panorpodes paradoxus</i>	Mecoptera			AF180061
<i>Panorpa claripennis</i>	Mecoptera			AF180068
<i>Neopanorpa</i> sp.	Mecoptera			AF180059
<i>Anopheles sinensis</i>	Diptera	Culicomorpha	Culicidae	AF373609
<i>Comptosia quadripennis</i>	Diptera	Asilomorpha	Bombyliidae	AJ391511
<i>Atrichochira</i> sp.	Diptera	Asilomorpha	Bombyliidae	AJ391528
<i>Villa fumea</i>	Diptera	Asilomorpha	Bombyliidae	AJ391530
<i>Megaselia</i> sp.	Diptera	Aschiza	Phoridae	AF154807
<i>Chalarus</i> sp.	Diptera	Aschiza	Pipunculidae	AF154757
<i>Jassidophaga</i> sp.	Diptera	Aschiza	Pipunculidae	AF154787
<i>Collinias</i> sp.	Diptera	Aschiza	Pipunculidae	AF154766
<i>Tomosvaryella</i> sp.	Diptera	Aschiza	Pipunculidae	AF154818
<i>Lindneromyia argentifascia</i>	Diptera	Aschiza	Platyppezidae	AF154810
<i>Ceriana</i> sp.	Diptera	Aschiza	Syrphidae	AF154755
<i>Eristalinus punctulatus</i>	Diptera	Aschiza	Syrphidae	AF154776
<i>Cyphipelta rufocyanea</i>	Diptera	Aschiza	Syrphidae	AF154772
<i>Toxomerus geminatus</i>	Diptera	Aschiza	Syrphidae	AF154819
<i>Chaetoceolopa sydneyensis</i>	Diptera	Acalyptratae	Coelopidae	AF403454
<i>Chaetoceolopa sydneyensis</i>	Diptera	Acalyptratae	Coelopidae	AF403454
<i>Dasycoelopa australis</i>	Diptera	Acalyptratae	Coelopidae	AF403465
<i>This canus</i>	Diptera	Acalyptratae	Coelopidae	AF403473
<i>Coelopa alluandi</i>	Diptera	Acalyptratae	Coelopidae	AF403455
<i>Neoconops longicornis</i>	Diptera	Acalyptratae	Conopidae	AF154769
<i>Helcomyza mirabilis</i>	Diptera	Acalyptratae	Helcomyzidae	AF403449
<i>Dryomyza anilis</i>	Diptera	Acalyptratae	Dryomyzidae	AF403448
<i>Mimegralla australica</i>	Diptera	Acalyptratae	Micropezidae	AF154794
<i>Orygma luctuosum</i>	Diptera	Acalyptratae	Sepsidae	AF403445
<i>Bactrocera tsuneonis</i>	Diptera	Acalyptratae	Tephritidae	AB074030
<i>Anastrepha ludens</i>	Diptera	Acalyptratae	Tephritidae	AB035102
<i>Bactrocera ishigakiensis</i>	Diptera	Acalyptratae	Tephritidae	AB035098
<i>Chrysomya putoria</i>	Diptera	Calyptratae	Calliphoridae	NC_002697
<i>Neosciomyza luteipennis</i>	Diptera	Calyptratae	Helosciomyzidae	AF403447
<i>Heterocheila buccata</i>	Diptera	Calyptratae	Heterocheilidae	AF403446



## Anexo IV

Táxons terminais utilizados na segunda análise do gene mitocondrial ribossomal 16S.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Suphisellus</i> sp.	Coleoptera			AY071792
<i>Pelodytes rotundatus</i>	Coleoptera			AY071790
<i>Lacconectus peguensis</i>	Coleoptera			AY071785
<i>Tenebrio molitor</i>	Coleoptera			AJ438153
<i>Carabus insulicola</i>	Coleoptera			AF219427
<i>Cicindela belfragei</i>	Coleoptera			AF133003
<i>Eubhyponomeutoides trachydeltus</i>	Lepidoptera			AY543628
<i>Anthanassa tulcis</i>	Lepidoptera			AF186920
<i>Helicoverpa armigera</i>	Lepidoptera			AB032068
<i>Neopsylla dispar</i>	Siphonaptera			AY337028
<i>Citellophilus sparsilis</i>	Siphonaptera			AY072638
<i>Rhadinopsylla daburica</i>	Siphonaptera			AY072637
<i>Amphipsylla quadratoides</i>	Siphonaptera			AY072636
<i>Culicoides imicola</i>	Diptera	Culicomorpha	Ceratopogonidae	AF281317
<i>Culicoides</i> sp.	Diptera	Culicomorpha	Ceratopogonidae	AY140904
<i>Culicoides tuttifrutti</i>	Diptera	Culicomorpha	Ceratopogonidae	AF281322
<i>Chironomus</i> sp.	Diptera	Culicomorpha	Chironomidae	AF456851
<i>Aedes aegypti</i>	Diptera	Culicomorpha	Culicidae	AF034475
<i>Anopheles nimpe</i>	Diptera	Culicomorpha	Culicidae	AF216284
<i>Culex bitaeniorhynchus</i>	Diptera	Culicomorpha	Culicidae	AF034474
<i>Austrosimulium bancrofti</i>	Diptera	Culicomorpha	Simuliidae	AB093130
<i>Prosimulium kiotoense</i>	Diptera	Culicomorpha	Simuliidae	AB056747
<i>Simulium birtinervis</i>	Diptera	Culicomorpha	Simuliidae	AB093111
<i>Chrysops carbonarius</i>	Diptera	Tabanomorpha	Tabanidae	AY123341
<i>Tabanus rufifrater</i>	Diptera	Tabanomorpha	Tabanidae	AY140887
<i>Dialysis</i> sp.	Diptera	Xylophagomorpha	Xylophagidae	AY140888
<i>Apiocera</i> sp.	Diptera	Asilomorpha	Apioceridae	AY325006
<i>Laphria thoracica</i>	Diptera	Asilomorpha	Asilidae	AY123342
<i>Stenopogon martini</i>	Diptera	Asilomorpha	Asilidae	AY325013
<i>Comptosia</i> sp.	Diptera	Asilomorpha	Bombyliidae	AF456847
<i>Villa</i> sp.	Diptera	Asilomorpha	Bombyliidae	AF456846
<i>Dolichopus</i> sp.	Diptera	Empidoidea	Dolichopodidae	AY123343
<i>Hilara maura</i>	Diptera	Empidoidea	Empididae	AF126346
<i>Mydas clavatus</i>	Diptera	Asilomorpha	Mydidae	AY324983
<i>Scenopinus fenestralis</i>	Diptera	Asilomorpha	Scenopinidae	AY325007
<i>Ectinorhynchus</i> sp.	Diptera	Asilomorpha	Therevidae	AF456856
<i>Ozodiceromyia costalis</i>	Diptera	Asilomorpha	Therevidae	AY325004
<i>Lonchoptera lutea</i>	Diptera	Aschiza	Lonchopteridae	AF126338
<i>Anevrina thoracica</i>	Diptera	Aschiza	Phoridae	AF126345
<i>Clitelloxenia andreyae</i>	Diptera	Aschiza	Phoridae	AF126339
<i>Conicera similis</i>	Diptera	Aschiza	Phoridae	AF126337
<i>Tripbleba distinguenda</i>	Diptera	Aschiza	Phoridae	AF126331
<i>Opetia nigra</i>	Diptera	Aschiza	Platyppezidae	AF126347
<i>Protocythia modesta</i>	Diptera	Aschiza	Platyppezidae	AF126348
<i>Eristalis tenax</i>	Diptera	Aschiza	Syrphidae	AY123344
<i>Melanostoma scalare</i>	Diptera	Aschiza	Syrphidae	AF126336
<i>Melanagromyza minimoides</i>	Diptera	Acalyptratae	Agromyzidae	AY573120
<i>Physocephala sagittaria</i>	Diptera	Acalyptratae	Conopidae	AY123349
<i>Chaetodiopsis meigenii</i>	Diptera	Acalyptratae	Diopsidae	AF304717
<i>Cyrtodiopsis currani</i>	Diptera	Acalyptratae	Diopsidae	AF304746

<i>Diopsis longicornis</i>	Diptera	Acalyptratae	Diopsidae	AF304745
<i>Amiota picta</i>	Diptera	Acalyptratae	Drosophilidae	AF479792
<i>Chymomyza amoena</i>	Diptera	Acalyptratae	Drosophilidae	M93986
<i>Stegana flavimana</i>	Diptera	Acalyptratae	Drosophilidae	AF479795

#### Anexo IV

#### Continuação

<i>Limmellia quadrata</i>	Diptera	Acalyptratae	Ephydriidae	AF126333
<i>Hydrellia modesta</i>	Diptera	Acalyptratae	Ephydriidae	AF126334
<i>Minettia longipennis</i>	Diptera	Acalyptratae	Lauxaniidae	AJ439184
<i>Sapromyza ultima</i>	Diptera	Acalyptratae	Lauxaniidae	AJ439209
<i>Sapromyza biscoitoi</i>	Diptera	Acalyptratae	Lauxaniidae	AJ439196
<i>Compsobata</i> sp.	Diptera	Acalyptratae	Micropezidae	AY573112
<i>Rainieria antennaeipes</i>	Diptera	Acalyptratae	Micropezidae	AY123347
<i>Delphinia picta</i>	Diptera	Acalyptratae	Otitidae	AF177121
<i>Tritoxa flexa</i>	Diptera	Acalyptratae	Otitidae	AY123351
<i>Rivellia quadrifasciata</i>	Diptera	Acalyptratae	Platystomatidae	AF177120
<i>Pyrgota undata</i>	Diptera	Acalyptratae	Pyrgotidae	AY123352
<i>Acanthonerva</i> sp.	Diptera	Acalyptratae	Tephritidae	AF177122
<i>Acidia cognate</i>	Diptera	Acalyptratae	Tephritidae	AF177141
<i>Stranzia vittigera</i>	Diptera	Acalyptratae	Tephritidae	AF177134
<i>Delia platura</i>	Diptera	Calypttratae	Anthomyiidae	AF086867
<i>Calliphora vomitoria</i>	Diptera	Calypttratae	Calliphoridae	AY114203
<i>Lucilia sericata</i>	Diptera	Calypttratae	Calliphoridae	AY114195
<i>Glossina morsitans</i>	Diptera	Calypttratae	Glossinidae	AF072373
<i>Lipoptena cervi</i>	Diptera	Calypttratae	Hippoboscidae	AF322437
<i>Musca domestica</i>	Diptera	Calypttratae	Muscidae	AY123346
Nycteribiidae sp.	Diptera	Calypttratae	Nycteribiidae	AF456855
<i>Penicillidia</i> sp.	Diptera	Calypttratae	Nycteribiidae	AF322435
<i>Penicillidia</i> sp.	Diptera	Calypttratae	Nycteribiidae	AF086866
<i>Cephenemyia stimulator</i>	Diptera	Calypttratae	Oestridae	AF322441
<i>Cuterebra</i> sp.	Diptera	Calypttratae	Oestridae	AF322440
<i>Gasterophilus haemorrhoidalis</i>	Diptera	Calypttratae	Oestridae	AF322439
<i>Hypoderma diana</i>	Diptera	Calypttratae	Oestridae	AF322438
<i>Rhinomorinia sarcophagina</i>	Diptera	Calypttratae	Rhinophoridae	AF322434
<i>Sarcophaga bullata</i>	Diptera	Calypttratae	Sarcophagidae	AF322433
<i>Archytas nivalis</i>	Diptera	Calypttratae	Tachinidae	AY123345
<i>Nemoraea pellucida</i>	Diptera	Calypttratae	Tachinidae	AF322432

## Anexo V

Táxons terminais utilizados na primeira análise do gene mitocondrial COI.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Cicindela punctulata</i>	Coleoptera			AY165729
<i>Hemicrepidius memnonius</i>	Coleoptera			AY165705
<i>Harmonia axyridis</i>	Coleoptera			AY165684
<i>Melitaea latonigena</i>	Lepidoptera			AY346234
<i>Tegeticula cassandra</i>	Lepidoptera			AY568273
<i>Kallima paralekta</i>	Lepidoptera			AY090229
<i>Agathipbaga queenslandensis</i>	Lepidoptera			AF436487
<i>Panorpa japonica</i>	Mecoptera			AF180087
<i>Panorpodes pulcher</i>	Mecoptera			AF180082
<i>Neopanorpa</i> sp.	Mecoptera			AF180079
<i>Panorpa meridionalis</i>	Mecoptera			AF180078
<i>Chrysoperla furcifera</i>	Neuroptera			AB081460
<i>Chrysoperla carnea</i>	Neuroptera			AB081459
<i>Helicopsyche borealis</i>	Trichoptera			AF436540
<i>Tipula</i> sp.	Diptera	Tipulomorpha	Tipulidae	AY165639
<i>Elporia barnardi</i>	Diptera	Blephariceromorpha	Blephariceridae	AF427038
<i>Asteralobia humuli</i>	Diptera	Bibionomorpha	Cecidomyiidae	AB176717
<i>Contarinia tritici</i>	Diptera	Bibionomorpha	Cecidomyiidae	AY485383
<i>Hartigiola faggalli</i>	Diptera	Bibionomorpha	Cecidomyiidae	AB162836
<i>Mikiola fagi</i>	Diptera	Bibionomorpha	Cecidomyiidae	AB162848
<i>Rhopalomyia struma</i>	Diptera	Bibionomorpha	Cecidomyiidae	AB162419
<i>Coboldia fuscipes</i>	Diptera	Bibionomorpha	Scatopsidae	AF320761
<i>Coboldia fuscipes</i>	Diptera	Bibionomorpha	Scatopsidae	AF319839
<i>Lycoriella mali</i>	Diptera	Bibionomorpha	Sciaridae	AF319838
<i>Bittacomorpha clavipes</i>	Diptera	Ptychopteromorpha	Ptychopteridae	AY165638
<i>Culicoides bolitinos</i>	Diptera	Culicomorpha	Ceratopogonidae	AF071931
<i>Culicoides circumscriptus</i>	Diptera	Culicomorpha	Ceratopogonidae	AF069248
<i>Anopheles sacharovi</i>	Diptera	Culicomorpha	Culicidae	AY135707
<i>Anopheles stephensi</i>	Diptera	Culicomorpha	Culicidae	AF181266
<i>Simulium venustum</i>	Diptera	Culicomorpha	Simuliidae	AY165663
<i>Atherix</i> sp.	Diptera	Tabanomorpha	Athericidae	AY165650
<i>Glutops rossi</i>	Diptera	Tabanomorpha	Pelecorhynchidae	AY165651
<i>Chrysopilus</i> sp.	Diptera	Tabanomorpha	Rhagionidae	AY165706
<i>Symphoromyia</i> sp.	Diptera	Tabanomorpha	Rhagionidae	AY165652
<i>Stratiomys</i> sp.	Diptera	Tabanomorpha	Stratiomyidae	AY165721
<i>Chrysops</i> sp.	Diptera	Tabanomorpha	Tabanidae	AY165640
<i>Macroceromyia</i> sp.	Diptera	Tabanomorpha	Xylomyidae	AY165711
<i>Xylophagus</i> sp.	Diptera	Xilophagomorpha	Xylophagidae	AY165704
<i>Asilus</i> sp.	Diptera	Asilomorpha	Asilidae	AY165713
<i>Efferia aestuans</i>	Diptera	Asilomorpha	Asilidae	AY165641
<i>Lasioptogon</i> sp.	Diptera	Asilomorpha	Asilidae	AY165732
<i>Anthrax</i> sp.	Diptera	Asilomorpha	Bombyliidae	AY165731
<i>Bombylius validus</i>	Diptera	Asilomorpha	Bombyliidae	AY165655
<i>Empis</i> sp.	Diptera	Asilomorpha	Empididae	AY165709
<i>Chrysotoxum</i> sp.	Diptera	Aschiza	Syrphidae	AY165714
<i>Scaeva pyrastris</i>	Diptera	Aschiza	Syrphidae	AY165733
<i>Hirtodrosophila histrioides</i>	Diptera	Acalyptatae	Drosophilidae	AF519387
<i>Suillia</i> sp.	Diptera	Acalyptatae	Heleomyzidae	AY165737
<i>Sphecomyiella valida</i>	Diptera	Acalyptatae	Pyrgotidae	AY165717
<i>Sphaerocera curvipes</i>	Diptera	Acalyptatae	Sphaeroceridae	AY165653

---

<i>Calliphora vomitoria</i>	Diptera	Calyptratae	Calliphoridae	AY092813
<i>Calliphora vicina</i>	Diptera	Calyptratae	Calliphoridae	AY092800
<i>Musca domestica</i>	Diptera	Calyptratae	Muscidae	AF259518
<i>Blaesoxipha plinthopyga</i>	Diptera	Calyptratae	Sarcophagidae	AF259514

---

## Anexo V

Continuação

---

<i>Wohlfahrtia vigil</i>	Diptera	Calyptratae	Sarcophagidae	AF259516
<i>Ravinia lberminieri</i>	Diptera	Calyptratae	Sarcophagidae	AF259513
<i>Sarcophaga africa</i>	Diptera	Calyptratae	Sarcophagidae	AF259508

---

## Anexo VI

Táxons terminais utilizados na segunda análise do gene mitocondrial COI.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Cicindela hybrida</i>	Coleoptera			AJ583540
<i>Ilybius chalconatus</i>	Coleoptera			AF428237
<i>Tenebrio molitor</i>	Coleoptera			AJ438095
<i>Erebia palarica</i>	Lepidoptera			AY480014
<i>Mantodea</i> sp.	Mantodea			AY318914
<i>Neopanorpa</i> sp.	Mecoptera			AF180098
<i>Panorpa communis</i>	Mecoptera			AF180090
<i>Panorpodes paradoxus</i>	Mecoptera			AF180100
<i>Chrysoperla carnea</i>	Neuroptera			AB081321
<i>Chrysoperla furcifera</i>	Neuroptera			AB081460
<i>Lype diversa</i>	Trichoptera			AF436582
<i>Parapsyche elsis</i>	Trichoptera			AF436568
<i>Chaoborus americanus</i>	Diptera	Culicomorpha	Chaoboridae	AJ427618
<i>Chironomus</i> sp.	Diptera	Culicomorpha	Chironomidae	AF456875
<i>Aedeomyia squamipennis</i>	Diptera	Culicomorpha	Culicidae	AF417729
<i>Aedes punctor</i>	Diptera	Culicomorpha	Culicidae	AF253035
<i>Anopheles marajoara</i>	Diptera	Culicomorpha	Culicidae	AF417699
<i>Bironella gracilis</i>	Diptera	Culicomorpha	Culicidae	AF417725
<i>Chagasia bathana</i>	Diptera	Culicomorpha	Culicidae	AF417726
<i>Toxorhynchites amboinensis</i>	Diptera	Culicomorpha	Culicidae	AF417727
<i>Uranotaenia lowii</i>	Diptera	Culicomorpha	Culicidae	AF417728
<i>Simulium excasperans</i>	Diptera	Culicomorpha	Simuliidae	AF333261
<i>Simulium malardei</i>	Diptera	Culicomorpha	Simuliidae	AF333266
<i>Comptosia</i> sp.	Diptera	Asilomorpha	Bombyliidae	AF456871
<i>Heteropsilopus ingenuus</i>	Diptera	Empidoidea	Dolichopodidae	AF456876
<i>Apocephalus paraponerae</i>	Diptera	Aschiza	Phoridae	AF217467
<i>Apocephalus</i> sp.	Diptera	Aschiza	Phoridae	AF217483
<i>Megaselia scalaris</i>	Diptera	Aschiza	Phoridae	AF217464
<i>Jassidophaga villosa</i>	Diptera	Aschiza	Pipunculidae	AY261685
<i>Agathomyia unicolor</i>	Diptera	Aschiza	Platypezidae	AY261684
<i>Alipumilio avispa</i>	Diptera	Aschiza	Syrphidae	AY261709
<i>Blera fallax</i>	Diptera	Aschiza	Syrphidae	AY261687
<i>Cheilosia convexifrons</i>	Diptera	Aschiza	Syrphidae	AY261691
<i>Paragus haemorrhous</i>	Diptera	Aschiza	Syrphidae	AY174466
<i>Xylota ignava</i>	Diptera	Aschiza	Syrphidae	AY212790
<i>Cerodontha fasciata</i>	Diptera	Acalyptratae	Agromyzidae	AF230241
<i>Liriomyza huidobrensis</i>	Diptera	Acalyptratae	Agromyzidae	AB106184
<i>Phytomyza jucunda</i>	Diptera	Acalyptratae	Agromyzidae	AF276840
<i>Anastrepha fraterculus</i>	Diptera	Acalyptratae	Tephritidae	AF420617
<i>Bactrocera oleae</i>	Diptera	Acalyptratae	Tephritidae	AY165957
<i>Chiastocheta curvibasis</i>	Diptera	Calyptratae	Anthomyiidae	AF325388
<i>Delia brassica</i>	Diptera	Calyptratae	Anthomyiidae	AF325362
<i>Lasiomma seminitidum</i>	Diptera	Calyptratae	Anthomyiidae	AF104624
<i>Calliphora vicina</i>	Diptera	Calyptratae	Calliphoridae	AJ417702
<i>Chrysomya putoria</i>	Diptera	Calyptratae	Calliphoridae	NC_002697
<i>Dyscritomyia fasciata</i>	Diptera	Calyptratae	Calliphoridae	AY074902
<i>Lucilia cuprina</i>	Diptera	Calyptratae	Calliphoridae	AY097335
<i>Fannia armata</i>	Diptera	Calyptratae	Fanniidae	AF104623
<i>Fergusonina purcelli</i>	Diptera	Calyptratae	Fergusoninidae	AY687981
<i>Fergusonina schefferae</i>	Diptera	Calyptratae	Fergusoninidae	AY687964

<i>Musca domestica</i>	Diptera	Calyptratae	Muscidae	AF104622
<i>Nycteribiidae</i> sp.	Diptera	Calyptratae	Nycteribiidae	AF456879
<i>Cephenemyia stimulator</i>	Diptera	Calyptratae	Oestridae	AF497768
<i>Cuterebra baeri</i>	Diptera	Calyptratae	Oestridae	AF497777

## Anexo VI

### Continuação

<i>Gasterophilus haemorrhoidalis</i>	Diptera	Calyptratae	Oestridae	AF497774
<i>Hypoderma bovis</i>	Diptera	Calyptratae	Oestridae	AF497761
<i>Acerocnema macrocera</i>	Diptera	Calyptratae	Scathophagidae	AF181025
<i>Chaetosa palpalis</i>	Diptera	Calyptratae	Scathophagidae	AF181017
<i>Cordilura albipes</i>	Diptera	Calyptratae	Scathophagidae	AF180995
<i>Nanna brunneicosta</i>	Diptera	Calyptratae	Scathophagidae	AF181006
<i>Spaziphora cincta</i>	Diptera	Calyptratae	Scathophagidae	AF181012
<i>Ormia lineifrons</i>	Diptera	Calyptratae	Tachinidae	AY860638
<i>Zelia vertebrata</i>	Diptera	Calyptratae	Tachinidae	AY864715

## Anexo VII

Táxons terminais utilizados na análise do gene mitocondrial COII.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Carabus splendens</i>	Coleoptera			AF175733
<i>Brachyderes rugatus calvus</i>	Coleoptera			AJ389860
<i>Opogona glycyphaga</i>	Lepidoptera			AY320506
<i>Hemileuca</i> sp.	Lepidoptera			AF423983
<i>Panorpodes pulcher</i>	Mecoptera			AF424039
<i>Panorpa communis</i>	Mecoptera			AF424024
<i>Nannochorista dipteroides</i>	Mecoptera			AF424019
<i>Bittacus pilicornis</i>	Mecoptera			AF423993
<i>Chrysoperla plorabunda</i>	Neuroptera			AF254097
<i>Parapsyllus magellanicus</i>	Siphonaptera			AF424042
<i>Pulex irritans</i>	Siphonaptera			AF424041
<i>Ctenocephalides felis</i>	Siphonaptera			M83964
<i>Oecetis avara</i>	Trichoptera			AF423986
<i>Pycnopsyche lepida</i>	Trichoptera			AF423984
<i>Epiphragma fasciapenne</i>	Diptera	Tipulomorpha	Limoniidae	AF423979
<i>Holorusia rubiginosa</i>	Diptera	Tipulomorpha	Tipulidae	AF423980
<i>Tanyptera dorsalis</i>	Diptera	Tipulomorpha	Tipulidae	AF423978
<i>Sylvicola fenestralis</i>	Diptera	Bibionomorpha	Anisopodidae	AF547667
<i>Cramptonomyia spenceri</i>	Diptera	Bibionomorpha	Pachyneuridae	AF547665
<i>Psychoda</i> sp.	Diptera	Psychodomorpha	Psychodidae	AF547666
<i>Ptychoptera</i> sp.	Diptera	Ptychopteromorpha	Ptychopteridae	AF547668
<i>Atrichopogon</i> sp.	Diptera	Culicomorpha	Ceratopogonidae	AF547674
<i>Ceratopogon annettae</i>	Diptera	Culicomorpha	Ceratopogonidae	AF547679
<i>Culicoides barbosai</i>	Diptera	Culicomorpha	Ceratopogonidae	AF547677
<i>Monobelea magnitbeca</i>	Diptera	Culicomorpha	Ceratopogonidae	AF547682
<i>Afrochilus harrisoni</i>	Diptera	Culicomorpha	Chironomidae	AY261448
<i>Austrochilus centralaustralis</i>	Diptera	Culicomorpha	Chironomidae	AY261450
<i>Chironomus calligraphus</i>	Diptera	Culicomorpha	Chironomidae	AJ310775
<i>Procladius villosimanus</i>	Diptera	Culicomorpha	Chironomidae	AY261453
<i>Aedeomyia squamipennis</i>	Diptera	Culicomorpha	Culicidae	AF417765
<i>Bironella gracilis</i>	Diptera	Culicomorpha	Culicidae	AF417761
<i>Culex pipiens quinquefasciatus</i>	Diptera	Culicomorpha	Culicidae	AF325716
<i>Ochlerotatus triseriatus</i>	Diptera	Culicomorpha	Culicidae	AF417766
<i>Toxorhynchites amboinensis</i>	Diptera	Culicomorpha	Culicidae	AF417763
<i>Cnephia</i> sp.	Diptera	Culicomorpha	Simuliidae	AF083859
<i>Helodon onychodactylus</i>	Diptera	Culicomorpha	Simuliidae	AF083857
<i>Prosimulium dicentum</i>	Diptera	Culicomorpha	Simuliidae	AF083858
<i>Simulium jacumbae</i>	Diptera	Culicomorpha	Simuliidae	AF083870
<i>Apiocera</i> sp.	Diptera	Asilomorpha	Apioceridae	AY325099
<i>Holopogon currani</i>	Diptera	Asilomorpha	Asilidae	AY325085
<i>Laphria</i> sp.	Diptera	Asilomorpha	Asilidae	AY325104
<i>Stenopogon martini</i>	Diptera	Asilomorpha	Asilidae	AY325106
<i>Comptosia</i> sp.	Diptera	Asilomorpha	Bombyliidae	AF484036
<i>Villa</i> sp.	Diptera	Asilomorpha	Bombyliidae	AF465477
<i>Heteropsilopus ingenuus</i>	Diptera	Empidoidea	Dolichopodidae	AF484041
<i>Mydas clavatus</i>	Diptera	Asilomorpha	Mydidae	AY325076
<i>Scenopinus fenestralis</i>	Diptera	Asilomorpha	Scenopinidae	AY325100
<i>Ozodiceromyia costalis</i>	Diptera	Asilomorpha	Therevidae	AY325097
<i>Cheilosia himantopus</i>	Diptera	Aschiza	Syrphidae	AY055208
<i>Cheilosia orthotricha</i>	Diptera	Aschiza	Syrphidae	AY055207



<i>Cerodontha fasciata</i>	Diptera	Acalyptratae	Agromyzidae	AF230241
<i>Liriomyza sativae</i>	Diptera	Acalyptratae	Agromyzidae	AF472350
<i>Phytomyza vomitoriae</i>	Diptera	Acalyptratae	Agromyzidae	AF276842
<i>Chaetocoelopa sydneyensis</i>	Diptera	Acalyptratae	Coelopidae	AF508520

## Anexo VII

### Continuação

<i>Coelopa frigida</i>	Diptera	Acalyptratae	Coelopidae	AF427653
<i>Dasycoelopa australis</i>	Diptera	Acalyptratae	Coelopidae	AF508526
<i>Gluma nitida</i>	Diptera	Acalyptratae	Coelopidae	AF508524
<i>Conops santaroi</i>	Diptera	Acalyptratae	Conopidae	AY573147
<i>Myopa buccata</i>	Diptera	Acalyptratae	Conopidae	AY573160
<i>Physocephala sagittaria</i>	Diptera	Acalyptratae	Conopidae	AY573166
<i>Chaetodiopsis meigenii</i>	Diptera	Acalyptratae	Diopsidae	AF304752
<i>Diopsis gnu</i>	Diptera	Acalyptratae	Diopsidae	AF304779
<i>Telogastrus entabeneensis</i>	Diptera	Acalyptratae	Diopsidae	AF304766
<i>Amiota picta</i>	Diptera	Acalyptratae	Drosophilidae	AF478412
<i>Chymomyza amoena</i>	Diptera	Acalyptratae	Drosophilidae	AF478415
<i>Scaptomyza albovittata</i>	Diptera	Acalyptratae	Drosophilidae	U94220
<i>Compsobata</i> sp.	Diptera	Acalyptratae	Micropezidae	AY573146
<i>Rainieria antennaepes</i>	Diptera	Acalyptratae	Micropezidae	AY573173
<i>Bactrocera cucurbitae</i>	Diptera	Acalyptratae	Tephritidae	AB093595
<i>Ceratitis capitata</i>	Diptera	Acalyptratae	Tephritidae	U53270
<i>Rhagoletis pomonella</i>	Diptera	Acalyptratae	Tephritidae	U53229
<i>Fucellia maritima</i>	Diptera	Calypttratae	Anthomyiidae	AF508527
<i>Calliphora vicina</i>	Diptera	Calypttratae	Calliphoridae	AJ417702
<i>Chrysomya putoria</i>	Diptera	Calypttratae	Calliphoridae	NC_002697
<i>Eucalliphora latifrons</i>	Diptera	Calypttratae	Calliphoridae	AF295557
<i>Onesia tibialis</i>	Diptera	Calypttratae	Calliphoridae	AY012569
<i>Haematobia irritans</i>	Diptera	Calypttratae	Muscidae	AY184818
<i>Hydrotaea aenescens</i>	Diptera	Calypttratae	Muscidae	AY184819
<i>Musca domestica</i>	Diptera	Calypttratae	Muscidae	AY184817
<i>Musca autumnalis</i>	Diptera	Calypttratae	Muscidae	AY184816
<i>Stomoxys calcitrans</i>	Diptera	Calypttratae	Muscidae	AY184815
Nycteribiidae sp.	Diptera	Calypttratae	Nycteribiidae	AF484044
<i>Gasterophilus</i> sp.	Diptera	Calypttratae	Oestridae	AF484042
<i>Hypoderma lineatum</i>	Diptera	Calypttratae	Oestridae	AF295558
<i>Archytas nivalis</i>	Diptera	Calypttratae	Tachinidae	AY573142

## Anexo VIII

Táxons terminais utilizados na análise do gene nuclear 18S. Os táxons marcados com \* foram utilizados na análise conjunta com o gene 28S, incluindo *missing data*.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Hyalophora cecropia</i>	Lepidoptera			AF443791
<i>Panorpa</i> sp.	Mecoptera			AF136857
<i>Ctenocephalides felis</i>	Siphonaptera			AF136859
<i>Epiphragma fasciapenne</i>	Diptera	Tipulomorpha	Limoniidae	AF286294
<i>Tipula paterifera</i>	Diptera	Tipulomorpha	Tipulidae	AF136855
<i>Lutzomyia shannoni</i>	Diptera	Psychodomorpha	Psychodidae	U48382
<i>Brumptomyia pintoii</i> *	Diptera	Psychodomorpha	Psychodidae	AJ391736
<i>Sycorax</i> sp. *	Diptera	Psychodomorpha	Psychodidae	AJ391747
<i>Dasyhelea</i> sp.	Diptera	Culicomorpha	Ceratopogonidae	AY257523
<i>Culicoides variipennis</i>	Diptera	Culicomorpha	Ceratopogonidae	U48380
<i>Coretbrella wirthi</i>	Diptera	Culicomorpha	Chaoboridae	U49736
<i>Eucorethra underwoodi</i>	Diptera	Culicomorpha	Chaoboridae	U07981
<i>Austrochilus centralaustralis</i>	Diptera	Culicomorpha	Chironomidae	AY257520
<i>Anopheles pseudopunctipennis</i>	Diptera	Culicomorpha	Culicidae	U49735
<i>Dixella cornuta</i>	Diptera	Culicomorpha	Dixidae	U48381
<i>Simulium sanctipanli</i>	Diptera	Culicomorpha	Simuliidae	AF403799
<i>Simulium vittatum</i> *	Diptera	Culicomorpha	Simuliidae	U48383
<i>Chrysops niger</i>	Diptera	Tabanomorpha	Tabanidae	AF073889
<i>Apiocera</i> sp. *	Diptera	Asilomorpha	Apioceridae	AY325037
<i>Laphria</i> sp.	Diptera	Asilomorpha	Asilidae	U65157
<i>Stenopogon martini</i>	Diptera	Asilomorpha	Asilidae	AY325044
<i>Hemipenthes jaennickeana</i>	Diptera	Asilomorpha	Bombyliidae	AY325036
<i>Mydas clavatus</i>	Diptera	Asilomorpha	Mydidae	AF423777
<i>Scenopinus fenestralis</i>	Diptera	Asilomorpha	Scenopinidae	AY325038
<i>Ozodiceromyia costalis</i>	Diptera	Asilomorpha	Therevidae	AY325035
<i>Liriomyza huidobrensis</i> *	Diptera	Acalyptatae	Agromyzidae	AY842478
<i>Drosophila melanogaster</i>	Diptera	Acalyptatae	Drosophilidae	X97143
<i>Lucilia</i> sp.	Diptera	Calyptatae	Calliphoridae	AF322425
<i>Glossina palpalis</i> *	Diptera	Calyptatae	Glossinidae	AF322431
<i>Lipoptena cervi</i> *	Diptera	Calyptatae	Hippoboscidae	AF322426
<i>Ornithomyia avicularia</i> *	Diptera	Calyptatae	Hippoboscidae	AF322421
<i>Ornithoica vicina</i> *	Diptera	Calyptatae	Hippoboscidae	AF073888
<i>Musca</i> sp.	Diptera	Calyptatae	Muscidae	AF322423
<i>Cephenemyia stimulator</i>	Diptera	Calyptatae	Oestridae	AF322429
<i>Cuterebra</i> sp.	Diptera	Calyptatae	Oestridae	AF322428
<i>Oestromyia leporina</i> *	Diptera	Calyptatae	Oestridae	AF322422
<i>Sarcophaga bullata</i>	Diptera	Calyptatae	Sarcophagidae	AF322419
<i>Nemoraea pellucida</i>	Diptera	Calyptatae	Tachinidae	AF322418

## Anexo IX

Táxons terminais utilizados na análise do gene nuclear 28S. Os táxons marcados com \* foram utilizados na análise conjunta com o gene 18S, incluindo *missing data*.

Espécie	Ordem	Subgrupo	Família	Nº de acesso (GenBank)
<i>Manduca sexta</i>	Lepidoptera			X93408
<i>Panorpa communis</i>	Mecoptera			X93406
<i>Ctenocephalides canis</i>	Siphonaptera			AF423974
<i>Hydropsyche</i> sp.	Trichoptera			AF338267
<i>Limonia nebulosa</i>	Diptera	Tipulomorpha	Limoniidae	X93378
<i>Tipula paludosa</i>	Diptera	Tipulomorpha	Tipulidae	X93405
<i>Sybicola fenestralis</i> *	Diptera	Bibionomorpha	Anisopodidae	X93376
<i>Dilophus febrilis</i> *	Diptera	Bibionomorpha	Bibionidae	X93375
<i>Clinodiplosis cilicrus</i> *	Diptera	Bibionomorpha	Cecidomyiidae	X93372
<i>Bolitophila cinerea</i> *	Diptera	Bibionomorpha	Mycetophilidae	X93373
<i>Anapausis inermis</i> *	Diptera	Bibionomorpha	Scatopsidae	X93374
<i>Bradysia coprophila</i> *	Diptera	Bibionomorpha	Sciaridae	X93402
<i>Psychoda cinerea</i>	Diptera	Psychodomorpha	Psychodidae	X93404
<i>Trichocera regelationis</i> *	Diptera	Psychodomorpha	Trichoceridae	X93379
<i>Culicoides punctatus</i>	Diptera	Culicomorpha	Ceratopogonidae	X89639
<i>Dasyhelea saxicola</i>	Diptera	Culicomorpha	Ceratopogonidae	X89640
<i>Coretbrella</i> sp.	Diptera	Culicomorpha	Chaoboridae	X89645
<i>Chaoborus crystallinus</i>	Diptera	Culicomorpha	Chaoboridae	X89644
<i>Chironomus tentans</i>	Diptera	Culicomorpha	Chironomidae	X93412
<i>Aedes albopictus</i>	Diptera	Culicomorpha	Culicidae	L22060
<i>Culex pipiens</i> *	Diptera	Culicomorpha	Culicidae	X93403
<i>Culex pipiens</i> *	Diptera	Culicomorpha	Culicidae	X89642
<i>Paradixa</i> sp.	Diptera	Culicomorpha	Dixidae	X89648
<i>Simulium euryadminiculum</i>	Diptera	Culicomorpha	Simuliidae	X93377
<i>Androprosopa larvata</i> *	Diptera	Culicomorpha	Thaumaleidae	X89647
<i>Chrysops carbonarius</i>	Diptera	Tabanomorpha	Tabanidae	AY123353
<i>Tabanus sudeticus</i> *	Diptera	Tabanomorpha	Tabanidae	X93371
<i>Apiocera</i> sp.	Diptera	Asilomorpha	Apioceridae	AY325068
<i>Laphria</i> sp.	Diptera	Asilomorpha	Asilidae	AY325073
<i>Stenopogon martini</i>	Diptera	Asilomorpha	Asilidae	AY325075
<i>Hemipenthes jaenickeana</i>	Diptera	Asilomorpha	Bombyliidae	AY325067
<i>Mydas clavatus</i>	Diptera	Asilomorpha	Mydidae	AY325045
<i>Scenopinus fenestralis</i>	Diptera	Asilomorpha	Scenopinidae	AY325069
<i>Ozodiceromyia costalis</i>	Diptera	Asilomorpha	Therevidae	AY325066
<i>Drosophila melanogaster</i>	Diptera	Acalyptratae	Drosophilidae	M21017
<i>Lucilia silvarum</i>	Diptera	Calypttratae	Calliphoridae	AJ551443
<i>Musca domestica</i>	Diptera	Calypttratae	Muscidae	AJ551427
<i>Gasterophilus intestinalis</i>	Diptera	Calypttratae	Oestridae	AJ551429
<i>Oestrus ovis</i>	Diptera	Calypttratae	Oestridae	AJ551428
<i>Sarcophaga carnaria</i>	Diptera	Calypttratae	Sarcophagidae	AJ551430
<i>Tachina grossa</i>	Diptera	Calypttratae	Tachinidae	AJ300130

## Anexo X

Táxons terminais utilizados na análise do gene nuclear 28S com os diferentes programas de alinhamento. Os terminais marcados com \* foram os terminais do grupo externo retirados na segunda etapa da análise e os marcados com \*\*, os terminais do grupo interno retirados das análises.

Espécie	Ordem	Subgrupo	Família	Número de acesso (GenBank)	
				D1	D7
<i>Manduca sexta</i>	Lepidoptera		Sphingidae	X93382	X93408
<i>Hydropsyche</i> sp. *	Trichoptera		Hydropsychidae		AF338267
<i>Panorpa communis</i>	Mecoptera		Panorpidae	X93380	X93406
<i>Panorpodes pulcher</i> *	Mecoptera		Panorpididae		AF423973
<i>Taeniochorista pallida</i> *	Mecoptera		Choristidae		AF423944
<i>Archaeopsylla erinacei</i>	Siphonaptera		Pulicidae	X93381	X93407
<i>Polygenis pradoi</i> *	Siphonaptera		Rhopalopsyllidae		AF423977
<i>Parapsyllus magellanicus</i> *	Siphonaptera		Rhopalopsyllidae		AF423976
<i>Pulex irritans</i> *	Siphonaptera		Pulicidae		AF423975
<i>Ctenocephalides canis</i> *	Siphonaptera		Pulicidae		AF423974
<i>Limonia nebulosa</i>	Diptera	Tipulomorpha	Limoniidae	X93369	X93378
<i>Tipula paludosa</i>	Diptera	Tipulomorpha	Tipulidae	X93387	X93405
<i>Culicoides punctatus</i> **	Diptera	Culicomorpha	Ceratopogonidae		X89639
<i>Dasybelea saxicola</i> **	Diptera	Culicomorpha	Ceratopogonidae		X89640
<i>Chaoborus crystallinus</i> **	Diptera	Culicomorpha	Chaoboridae		X89644
<i>Corethrella</i> sp. **	Diptera	Culicomorpha	Corethrellidae		X89645
<i>Chironomus tentans</i>	Diptera	Culicomorpha	Chironomidae	X93383	X93412
<i>Aedes albopictus</i>	Diptera	Culicomorpha	Culicidae		L22060
<i>Culex pipiens</i>	Diptera	Culicomorpha	Culicidae	X93384	X93403
<i>Culex pipiens</i> **	Diptera	Culicomorpha	Culicidae		X89642
<i>Paradixa</i> sp. **	Diptera	Culicomorpha	Dixidae		X89648
<i>Simulium euryadminiculum</i>	Diptera	Culicomorpha	Simuliidae	X93368	X93377
<i>Androprosopa larvata</i> **	Diptera	Culicomorpha	Thaumaleidae		X89647
<i>Bolitophila cinerea</i>	Diptera	Bibionomorpha	Mycetophilidae	X93365	X93373
<i>Bradysia coprophila</i>	Diptera	Bibionomorpha	Sciaridae	X93391	X93402
<i>Syhricola fenestralis</i>	Diptera	Bibionomorpha	Anisopodidae	X93367	X93376
<i>Dilophus febrilis</i>	Diptera	Bibionomorpha	Bibionidae	X93366	X93375
<i>Clinodiplosis cilicrus</i>	Diptera	Bibionomorpha	Cecidomyiidae	X93363	X93372
<i>Anapausis inermis</i>	Diptera	Psychodomorpha	Scatopsidae	X93364	X93374
<i>Psychoda cinerea</i>	Diptera	Psychodomorpha	Psychodidae	X93386	X93404
<i>Trichocera regelationis</i>	Diptera	Psychodomorpha	Trichoceridae	X93370	X93379
<i>Drosophila melanogaster</i>	Diptera	Brachycera	Drosophilidae		M21017
<i>Tabanus sudeticus</i>	Diptera	Brachycera	Tabanidae	X93362	X93371

## Anexo XI

```
Macro=;
var =
    0 active_pos
    1 run_length
    2 first_len
    3 sec_len
    4 third_len
    5 tot_len
    6 ild_value
    7 tv_cost
    8 ts_cost
    9 redundancy;
log ild_table.txt;
lquote];
quote 'TV &09 TS &09 1a Posit &09 2a Posit &09 3a Posit &09 Total &09 IDL;
log/;
log ana.log
/* -- unweighted run -- */
ccode;
costs;
hold 10000; mult=replic 10; keep 5; ratchet=iter 20; drift=iter 20; tchoose/; length;
set tot_len length[0];
loop 1 3
cc ]0.;
set active_pos #1; /* use 1 2 3 para ativar a posição */
quote ----- Active position 'active_pos';
    if ('active_pos'==1)
        loop 0+3 nchar /* ativa a primeira posição */
        if (#2==0)
            cc [0;
        else
            cc [ #2;
        end
    stop;
end
    if ('active_pos'==2)
        loop 1+3 nchar /* ativa a segunda posição */
        if (#2==1)
            cc [1;
        else
            cc [ #2;
        end
    stop;
end
    if ('active_pos'== 3)
        loop 2+3 nchar /* ativa a terceira posição */
```

```

        if (#2==2)
            cc [2;
            else
            cc [ #2;
            end
        stop;
    end
    ccode;
    costs;
    hold 10000; mult=replic 10; keep 5; ratchet=iter 20; drift=iter 20; tchoose/; length;
    set run_length length[0];
        if ('active_pos'==1)
            set first_len length[0];
            end
        if ('active_pos'==2)
            set sec_len length[0];
            end
        if ('active_pos'== 3)
            set third_len length[0];
            end
    stop
        set ild_value (('tot_len'-'(first_len'+ 'sec_len'+ 'third_len'))/'tot_len');
log+ ild_table.txt;
lquote];
quote 1 &09 1 &09 'first_len' &09 'sec_len' &09 'third_len' &09 'tot_len' &09 'ild_value';
log/;
log+ ana.log;

/* ----- */

loop 2+2 4
loop 0+2 4
cc [0;
if (#1==#2)
    continue
else
set tv_cost #1;
set ts_cost #2;
quote -----> Cost functions used: TV = #1 and TS = #2.;
smatrix=0 (firstset) A/C 'tv_cost' A/G 'ts_cost' A/T 'tv_cost' C/G 'tv_cost' C/T 'ts_cost'
G/T 'tv_cost'; ccode ( 0.; smatrix+firstset 0.;
ccode;
costs;
hold 10000; mult=replic 10; keep 5; ratchet=iter 20; drift=iter 20; tchoose/; length;
set tot_len length[0];
loop 1 3
cc ]0.;

```

```

set active_pos #3; /* use 1 2 3 para ativar a posição */
quote ----- Active position 'active_pos';
  if ('active_pos'==1)
    loop 0+3 nchar /* ativa a primeira posição */
    if (#4==0)
      cc [0;
    else
      cc [ #4;
    end
    stop;
  end
  if ('active_pos'==2)
    loop 1+3 nchar /* ativa a segunda posição */
    if (#4==1)
      cc [1;
    else
      cc [ #4;
    end
    stop;
  end
  if ('active_pos'== 3)
    loop 2+3 nchar /* ativa a terceira posição */
    if (#4==2)
      cc [2;
    else
      cc [ #4;
    end
    stop;
  end
  end
  ccode;
  costs;
  smatrix=0 (firstset) A/C 'tv_cost' A/G 'ts_cost' A/T 'tv_cost' C/G 'tv_cost' C/T
'ts_cost' G/T 'tv_cost'; ccode ( 0.; smatrix+firstset 0.;
  hold 10000; mult=replic 50; keep 5; ratchet=iter 20; drift=iter 20; tchoose/; length;
  set run_length length[0];
    if ('active_pos'==1)
      set first_len length[0];
    end
    if ('active_pos'==2)
      set sec_len length[0];
    end
    if ('active_pos'== 3)
      set third_len length[0];
    end
    set ild_value (((tot_len-'first_len'+ 'sec_len'+ 'third_len'))/'tot_len'));
  stop
log/;
log+ ild_table.txt;

```

```
lquote];  
quote 'tv_cost' &09 'ts_cost' &09 'first_len' &09 'sec_len' &09 'third_len' &09 'tot_len' &09  
'ild_value';  
log/;  
log+ ana.log;  
end;  
stop  
stop  
log/;  
end;  
proc/;
```